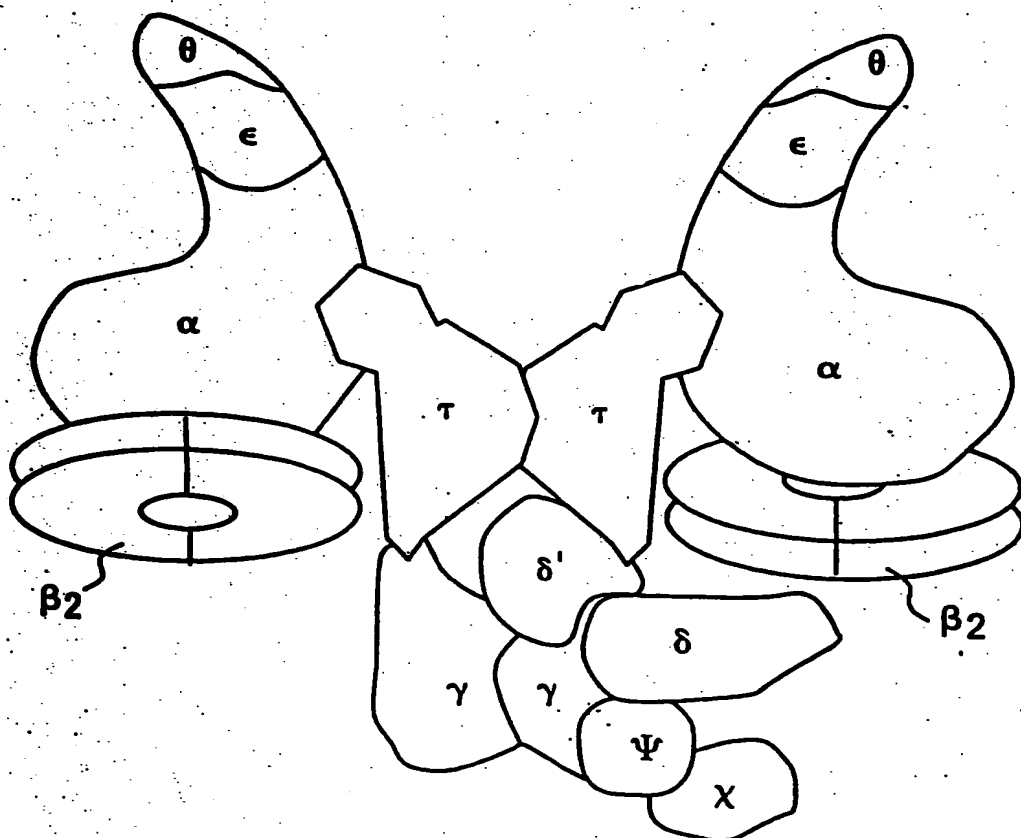


FIG.1



# ATP binding

E. coli  
 MSYQVLARKWRPQTFADVVGQEHVLTALANGLSLGRIHHAYLFSGTRGVGKTSIARLLAK  
 B. subtilis  
 MSYQALYRVFRPQRFEDVVGQEHITKTLQNALLOKKFSHAYLFSGPRGTGKTSAAKIFAK  
 \*\*\* \*

E. coli  
 GLNCETGITATPCGVCDNCREIEQGRFVDLIEIDAASRTKVEDTRDLLDNVQYAPARGRF  
 B. subtilis  
 AVNCEHAPVDEPCNECAACKGITNGSISDVIEIDAASNNGVDEIRDIRDVKFAPSAVTY  
 \*\*\* \*

E. coli  
 KVYLIDEVHMLSRHSFNALLKTLLEPPPEHVKFLLATTPQKLPVTILSRCLQFHLKALDV  
 B. subtilis  
 KVIIDEVHMLSIGAFNALLKTLLEPPPEHCIFILATTEPHKIPLTIISRCQRFDFKRITS  
 \*\*\* \*

FIG. 2

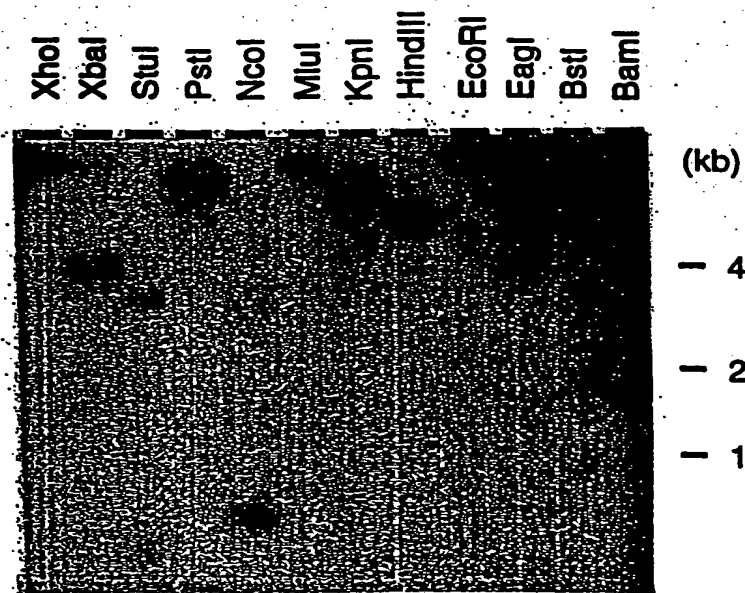


FIG.3

|   |   |            |             |             |                       |       |
|---|---|------------|-------------|-------------|-----------------------|-------|
| TCCGGGGGTG  | GGGTCCCG  | GTAGACCCG  | GCCCCCCTCCG | TGAGCCCCCTT | TACCCAGGCC            | 60    |
| GCCACCTCCT  | CCAGGGGGGC  | CAAGCGGTGC | AAGGAGAGGA  | ACGTCCGCAC  | <del>CACGCCCTAT</del> | 120   |
| ACTAGCCTT   | GTG AGC GCC CTC TAC CGC CGC TTC CGC CCC CTC ACC TTC CAG GAG GTG GTG |            |             |             | <sup>s.D.</sup>       | 180   |
|   | met ser ala leu tyr arg phe arg pro leu thr phe gln glu val val     |            |             |             |                       | (17)  |
| GGG CAG GAG CAC GTG AAG GAG CCC CTC AAG GCC ATC CGG GAG GGG AGG CTC GCC CAG     |   |            |             |             | CAC                   | 240   |
| gly gln glu his val lys glu pro leu lys ala ile arg glu arg leu ala gln         |   |            |             |             |                       | (37)  |
| GCS TAC CTC TTC TCC GGS AC  |   |            |             |             |                       |       |
| GCC TAC CTC TTC TCC GGG CCC AGG GGC GTG GGC AAG ACC ACG GCG AGG CTC CTC GCC     |   |            |             |             |                       | 300   |
| ala tyr leu phe ser gly pro arg gly val gly lys thr thr ala arg leu ala         |   |            |             |             |                       | (57)  |
| ATG GCG GTG GGG TGC CAG GGG GAA GAC CCC CCT TGC GGG GTC TGC CCC CAC TGC CAG GCG |   |            |             |             |                       | 360   |
| met ala val gly cys gln gly glu asp pro pro cys gly val cys pro his cys gln ala |   |            |             |             |                       | (77)  |
| GtG CAG AGG GGC GCC CAC CCG GAC GTG GTG GAC ATT GAC GCC GCC AGC AAC AAC TCC GTG |   |            |             |             |                       | 420   |
| val gln arg gly ala his pro asp val val asp ile asp ala ala ser asn ser val     |   |            |             |             |                       | (97)  |
| GAG GAC GTG CGG GAG CTG AGG GAA AGG ATC CAC CTC GCC CTC TCT GCC CCC AGG AAG     |   |            |             |             |                       | 480   |
| glu asp val arg glu leu arg glu arg ile his leu ala pro leu ser ala pro arg lys |   |            |             |             |                       | (117) |
| GTC TTC ATC CTG GAC GAG GCC CAC ATG CTC TCC AAA AGC GCC TTC AAC GCC CTC CTC AAG |   |            |             |             | <sup>C</sup>          | 540   |
| val phe ile leu asp Glu ala his met leu ser lys ser ala phe asn ala leu leu lys |   |            |             |             |                       | (137) |

FIG.4A-1



|   |       |
|---|-------|
| GAG CGC CTC GCC CGC CGC TCC GAC GCC TTA AGC CTG GAG GTG GCC CTC CTG GAG GCG GGA | 1140  |
| glu arg leu ala arg arg ser asp ala leu ser leu glu val ala leu leu glu ala gly | (337) |
| AGG GCC CTG GCC GAG GCC CTA CCC CAG CCC ACC GGC GCT CCT TCC CCA GAG GTC GGC     | 1200  |
| arg ala leu ala ala glu ala leu pro gln pro thr gly ala pro ser pro glu val gly | (357) |
| CCC AAG CCG GAA AGC CCC CCG ACC CCG GAA CCC CCA AGG CCC GAG GAG CCC GAC CTG     | 1260  |
| pro lys pro glu ser pro pro thr pro glu pro arg pro glu ala pro asp leu         | (377) |
| CGG GAG CGG TGG CGG GCC TTC CTC GAG GCC CTC AGG CCC ACC CTA CGG GCC TTC GTG CGG | 1320  |
| arg glu arg trp arg ala phe leu glu ala leu arg pro thr leu arg ala phe val arg | (397) |
| GAG GCC CGC CGC GAG GTC CCG GAA GGC CAG CTC TGC CTC GCT TTC CCC GAG GAC AAG GCC | 1380  |
| glu ala arg pro glu val arg glu gly gln leu cys leu ala phe pro glu asp lys ala | (417) |
| TTC CAC TAC CGC AAG GCC TCG GAA CAG AAG GTG AGG CTC CTC CCC CTG GCC CAG GCC CAT | 1440  |
| phe his tyr arg lys ala ser glu gln lys val arg leu leu pro leu ala gln ala his | (437) |
| frameshift site   |       |
| TTC GGG GTG GAG GAG GTC CTC CTC GAG GGA GAA AAA AAA AGC CTG AGC CCA AGG         | 1500  |
| phe gly val glu glu val leu val leu glu gly glu lys lys ser leu ser pro arg     | (457) |

FIG.4B-1

|   |       |
|---|-------|
| CCC CGC CCG GCC CCA CCT CCT GAA GCG CCC GCA CCC CCG GGC CCT CCC GAG GAG GAG GTA | 1560  |
| pro arg pro ala pro pro pro glu ala pro ala pro pro gly pro pro glu glu val     | (477) |
| GAG GCG GAG GAA GCG GCG GAG GAG GCC CCG GAG GAG GCC TTG AGG CCG GTG GTC CGC CTC | 1620  |
| glu ala glu glu ala ala glu glu ala pro glu ala leu arg arg val val arg leu     | (497) |
| CTG GGG GGG CCG GTG CTC TGG GTG CCG CCG CCC AGG ACC CCG GAG GCG CCG GAG GAG GAA | 1680  |
| leu gly gly arg val leu trp val arg arg pro arg thr arg glu ala pro glu glu     | (517) |
| CCC CTG AGC CAA GAC GAG ATA GGG GGT ACT GGT ATA TAA                             | 1740  |
| pro leu ser gln asp glu ile gly thr gly ile *                                   | (529) |
| CGACCTCGGA CAAGAGACCG TGGACAACAT CCTCAAGCGC CTCCGCCGTA TTGAGGGCCA               | 1820  |
| GGTGC GG GGG CTCCAGAAGA TGGTGGCCGA GGGCCGCCCC TCGGACGAGG TCCTCACCCA             | 1880  |
| GATGACCGCC ACCAAGAAGG CCATGGAGGC GCGGCCACC CTGATCCTCC ACGAGTTCCT                | 1940  |
| GAACGTCTGC GCCGCCGAGG TCTCCGAGGG CAAGTGAAC CCCAAGAAGC CCGAGGAGAT                | 2000  |
| CGCCACCATG CTGAAGAAGT TCATCTA   | 2027  |

FIG.4B-2

|             |         |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-------------|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| GGG CAG GAG | GTG AGC | GCC | CTC | TAC | CGC | CGC | TTC | CGC | CCC | CTC | ACC | TTC | CAG | GAG | GTG | GTG | 51   |
| GCC TAC CTC | TCC     | GGG | CCC | AGG | GGC | ATC | AAG | GCC | ATC | CGG | GAG | GGG | AGG | CTC | GCC | CAG | 111  |
| ATG GCG GTG | TGC     | CAG | GGG | GAA | GAC | CCC | CCT | TGC | GGG | GTC | ACC | ACG | AGG | CTC | CTC | GCC | 171  |
| GtG CAG AGG | GGC     | CAC | CCG | GAC | GTG | ATT | GAC | ATC | GAC | GCC | GCC | TGC | CCC | CAC | TGC | CAG | 231  |
| GAG GAC GTG | CGG     | GAG | CTG | AGG | ATC | AGG | ATC | AGG | ATC | GCC | CTC | AGC | AAC | AAC | TCC | GTG | 291  |
| GTC TTC ATC | CTG     | GAC | GCC | CAC | ATG | CTC | TCC | AAA | AGC | GCC | TTC | AAC | GCC | CTC | CTC | AAG | 351  |
| ACC CTG GAG | CCC     | CCC | CCC | CAC | GTC | CTC | TTC | GTC | TTC | GCC | ACC | ACC | GAG | CCC | GAG | AGG | 411  |
| ATG CCC ACC | ATC     | CTC | TCC | CGC | ACC | CGC | CAC | TTC | CGC | TTC | CGC | CGC | CTC | ACG | GAG | GAG | 471  |
| GAG ATC GCC | TTT     | AAG | CTC | CGC | ATC | CGC | GAG | GCC | GTG | GGG | CGG | GAG | GCG | GAG | GAG | GAG | 531  |
| GCC CTC CTC | CTC     | GCC | CGC | CTG | GCG | GAC | GGG | GCC | CTT | AGG | GAC | GCG | GAA | AGC | CTC | CTG | 591  |
| GAG CGC TTC | CTC     | CTC | GAA | GGC | CCC | CTC | ACC | CGG | AAG | GAG | GTG | GAG | CGC | GCC | CTA | GGC | 651  |
| TCC CCC CCA | GGG     | ACC | GGG | GTG | GCC | ATC | GCC | GCC | TCC | CTC | GCG | AGG | GGG | AAA | ACG | GCG | 711  |
| GAG GCC CTG | GGC     | CTC | CGG | CGC | CTC | TAC | GGG | GAA | GGG | TAC | GCC | CGC | AGG | AGC | CTG | GTC | 771  |
| TCG GGC CTT | TTG     | GAG | GTG | TTC | CGG | GAA | GCC | CTC | TAC | GCC | TTC | GGC | CTC | GCG | GGA | ACC | 831  |
| CCC CTT CCC | GCC     | CCG | CCC | CAG | GCC | CTG | ATC | GCC | ATG | ACC | GCC | CTG | GAC | GAG | GCC | ATG | 891  |
| GAG CGC CTC | GCC     | CGC | TCC | GAC | GCC | TTA | AGC | CTG | GAG | GTG | GCC | CTC | CTG | GAG | GCG | GGA | 951  |
| AGG GCC CTG | GCC     | GCC | GAG | GCC | CTA | CCC | CAG | CCC | ACG | GGC | GCT | CCT | TCC | CCA | GAG | GTC | 1011 |
| CCC AAG CCG | GAA     | AGC | CCC | ACC | ACC | CCG | GAA | CCC | CCA | AGG | CCC | GAG | GCG | CCC | GAC | CTG | 1071 |
| CGG GAG CCG | TGG     | CGG | GCC | TTC | CTC | GAG | GCC | CTC | AGG | CCC | ACC | CTA | CGG | GCC | TTC | CGG | 1131 |
| GAG GCC CCG | CGC     | GAG | GTG | CGG | GAA | GGC | CAG | CTC | TGC | CTC | GCT | TTC | CCC | GAG | AAG | GCC | 1191 |
| TTC CAC TAC | CGC     | AAG | GCC | TCG | GAA | CAG | AAG | GTG | AGG | CTC | CCC | CTG | GCC | CAG | GCC | CAT | 1251 |
| TTC GGG GTG | GAG     | GAG | GTG | GTC | CTC | GTC | CTG | GAG | GGA | GAA | AAA | AGC | CTG | AGC | CCA | AGG | 1311 |
| CCC CGC CCG | GCC     | CCA | CCT | CCT | GAA | GCG | CCC | GCA | CCC | CGG | CCT | CCC | GAG | GAG | GTA | GTA | 1371 |
| GAG GCG GAG | GAA     | GCG | GAG | GAG | GCC | CGG | GAG | GAG | GCC | TTG | AGG | CGG | GTG | GTC | CGC | CTC | 1431 |
| CTG GGG GGG | CGG     | GTG | CTC | TGG | GTG | CGG | CCC | AGG | ACC | CGG | GAG | GCG | CCG | GAG | GAG | GAA | 1491 |
|             |         |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 1551 |

CCC CTG AGC CAA GAC GAG ATA GGG GGT ACT GGT ATA TAA (1590)

FIG.4C



|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | ser | ala | leu | tyr | arg | arg | phe | leu | thr | phe | gln | gln | val | gly | gln | gln | 20  |
| his | val | lys | glu | pro | leu | lys | ala | ile | arg | glu | gly | arg | leu | ala | gln | ala | 40  |
| phe | ser | gly | pro | arg | gly | val | gly | thr | thr | thr | ala | arg | leu | ala | met | ala | 60  |
| gly | cys | gln | gly | glu | asp | pro | pro | cys | gly | val | cys | pro | his | cys | gln | ala | 80  |
| gly | ala | his | pro | asp | val | val | asp | ile | asp | ala | ala | ser | asn | ser | val | gln | 100 |
| arg | glu | leu | arg | glu | arg | ile | his | leu | ala | pro | leu | ser | ala | pro | arg | lys | 120 |
| leu | asp | glu | ala | his | met | leu | ser | lys | ser | ala | phe | asn | ala | leu | leu | lys | 140 |
| glu | pro | pro | pro | his | val | leu | phe | val | phe | ala | thr | thr | glu | pro | glu | arg | 160 |
| thr | ile | leu | ser | arg | thr | gln | his | phe | arg | phe | arg | arg | leu | thr | glu | gln | 180 |
| phe | lys | leu | arg | arg | ile | leu | glu | ala | val | gly | arg | glu | ala | glu | glu | ala | 200 |
| leu | leu | ala | arg | leu | ala | asp | gly | ala | leu | arg | asp | ala | glu | ser | leu | leu | 220 |
| leu | thr | gly | val | ala | gly | pro | leu | thr | arg | lys | glu | val | glu | ala | leu | gly | 240 |
| gly | thr | leu | ala | arg | leu | ile | ala | ala | ser | leu | ala | arg | gly | lys | thr | ala | 260 |
| gly | leu | ala | arg | arg | leu | tyr | gly | gly | tyr | ala | pro | arg | ser | leu | val | ser | 280 |
| leu | glu | val | phe | arg | glu | gly | leu | tyr | ala | ala | phe | gly | leu | ala | gly | thr | 300 |
| ala | pro | pro | gln | ala | leu | ile | ala | met | thr | ala | leu | asp | glu | ala | met | glu | 320 |
| ala | arg | arg | ser | asp | ala | leu | ser | leu | glu | val | ala | leu | leu | glu | ala | gly | 340 |
| ala | ala | glu | ala | leu | pro | gln | pro | thr | gly | ala | pro | ser | pro | glu | val | gly | 360 |
| glu | ser | pro | pro | thr | pro | glu | pro | pro | arg | pro | glu | ala | pro | asp | leu | arg | 380 |
| trp | arg | ala | phe | leu | glu | ala | leu | ala | thr | leu | arg | ala | phe | val | arg | glu | 400 |
| pro | glu | val | arg | glu | gly | gln | lys | leu | ala | phe | pro | glu | asp | lys | ala | phe | 420 |
| arg | lys | ala | ser | glu | gln | lys | val | arg | leu | pro | leu | ala | gln | ala | his | phe | 440 |
| glu | glu | val | val | leu | val | leu | glu | gly | lys | lys | ser | leu | ser | pro | arg | pro | 460 |
| ala | pro | pro | pro | glu | ala | pro | ala | pro | gly | pro | pro | glu | glu | val | glu | ala | 480 |
| glu | ala | ala | glu | glu | ala | pro | glu | ala | leu | arg | arg | val | val | arg | leu | leu | 500 |
| arg | val | leu | trp | val | arg | arg | pro | arg | thr | arg | glu | ala | pro | glu | glu | pro | 520 |
| gln | asp | glu | ile | gly | gly | thr | gly | thr | ile |     |     |     |     |     |     |     | 529 |

FIG.4D

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | ser | ala | leu | tyr | arg | arg | phe | arg | pro | leu | thr | phe | gln | glu | val | val | gly | gln | glu | 20  |     |
| his | val | lys | glu | pro | leu | lys | ala | ile | arg | glu | thr | gly | arg | leu | ala | gln | ala | tyr | leu | 40  |     |
| phe | ser | gly | pro | arg | gly | val | gly | lys | thr | thr | thr | ala | arg | leu | ala | met | ala | val |     | 60  |     |
| gly | cys | gln | gly | glu | asp | pro | pro | cys | gly | val | cys | pro | his | cys | gln | ala | val | gln | arg | 80  |     |
| gly | ala | his | pro | asp | val | val | asp | ile | asp | ala | ala | ser | asn | asn | ser | val | glu | asp | val | 100 |     |
| arg | glu | leu | arg | glu | arg | ile | his | leu | ala | pro | leu | ser | ala | pro | arg | lys | val | phe | ile | 120 |     |
| leu | asp | glu | ala | his | met | leu | ser | lys | ser | ala | phe | asn | ala | leu | leu | lys | thr | leu | glu | 140 |     |
| glu | pro | pro | pro | his | val | leu | phe | val | phe | ala | thr | thr | glu | pro | glu | arg | met | pro | pro | 160 |     |
| thr | ile | leu | ser | arg | thr | gln | his | phe | arg | phe | arg | arg | leu | thr | glu | glu | glu | ile | ala | 180 |     |
| phe | lys | leu | arg | arg | ile | leu | glu | ala | val | gly | arg | glu | ala | glu | glu | ala | leu | leu |     | 200 |     |
| leu | leu | ala | arg | leu | ala | asp | gly | ala | leu | arg | asp | ala | glu | ser | leu | leu | glu | arg | phe | 220 |     |
| leu | leu | leu | glu | gly | pro | leu | thr | arg | lys | glu | val | glu | arg | ala | leu | gly | ser | pro | pro | 240 |     |
| gly | thr | gly | val | ala | glu | ile | ala | ala | ser | leu | ala | arg | gly | lys | thr | ala | glu | ala | leu | 260 |     |
| gly | leu | ala | arg | arg | leu | tyr | gly | glu | gly | tyr | ala | pro | arg | ser | leu | val | ser | gly | leu | 280 |     |
| leu | glu | val | phe | arg | glu | gly | leu | tyr | ala | ala | phe | gly | leu | ala | gly | thr | pro | leu | pro | 300 |     |
| ala | pro | pro | gln | ala | leu | ile | ala | ala | met | thr | ala | leu | asp | glu | ala | met | glu | arg | leu | 320 |     |
| ala | arg | arg | ser | asp | ala | leu | ser | leu | glu | val | ala | leu | leu | glu | ala | gly | arg | ala | leu | 340 |     |
| ala | ala | glu | ala | leu | pro | gln | pro | thr | gly | ala | pro | ser | pro | glu | val | gly | pro | lys | pro | 360 |     |
| glu | ser | pro | pro | thr | pro | pro | glu | pro | arg | pro | glu | glu | ala | pro | asp | leu | arg | glu | arg | 380 |     |
| trp | arg | ala | phe | leu | glu | ala | ala | leu | arg | pro | thr | leu | arg | ala | phe | val | arg | glu | ala | arg | 400 |
| pro | glu | val | arg | glu | gly | gln | leu | leu | cys | leu | phe | pro | glu | asp | lys | ala | phe | his | tyr | 420 |     |
| arg | lys | ala | ser | glu | gln | lys | val | arg | leu | leu | pro | leu | ala | gln | ala | his | phe | gly | val | 440 |     |
| glu | glu | val | val | leu | val | leu | glu | gly | glu | lys | lys | pro | asp | pro | lys | ala | pro | pro | pro | 460 |     |
| gly | pro | thr | ser |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 464 |     |

FIG.4E

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | ser | ala | leu | tyr | arg | arg | phe | arg | pro | leu | thr | phe | gln | glu | val | val | gly | gln | glu | 20  |
| his | val | lys | glu | pro | arg | leu | lys | ala | ile | arg | glu | gly | arg | leu | ala | gln | ala | tyr | leu | 40  |
| phe | ser | gly | pro | arg | gly | val | gly | lys | thr | thr | thr | ala | arg | leu | ala | met | ala | val |     | 60  |
| gly | cys | gln | gly | glu | asp | pro | pro | cys | gly | val | cys | pro | his | cys | gln | ala | val | gln | arg | 80  |
| gly | ala | his | pro | asp | val | val | asp | ile | asp | ala | ala | ser | asn | asn | ser | val | glu | asp | val | 100 |
| arg | glu | leu | arg | glu | arg | ile | his | leu | ala | pro | leu | ser | ala | pro | arg | lys | val | phe | ile | 120 |
| leu | asp | glu | ala | his | met | leu | ser | lys | ser | ala | phe | asn | ala | leu | leu | lys | thr | leu | glu | 140 |
| glu | pro | pro | pro | his | val | leu | phe | val | phe | ala | thr | thr | glu | pro | glu | arg | met | pro | pro | 160 |
| thr | ile | leu | ser | arg | thr | thr | gln | his | phe | arg | phe | arg | leu | thr | glu | glu | glu | ile | ala | 180 |
| phe | lys | leu | arg | arg | ile | leu | glu | ala | val | gly | arg | glu | ala | glu | glu | glu | ala | leu | leu | 200 |
| leu | leu | ala | arg | leu | ala | asp | gly | ala | leu | arg | asp | ala | glu | ser | leu | leu | glu | arg | phe | 220 |
| leu | leu | leu | glu | gly | pro | leu | thr | arg | lys | glu | val | glu | arg | ala | leu | gly | ser | pro | pro | 240 |
| gly | thr | gly | val | ala | glu | ile | ala | ala | ser | leu | ala | arg | gly | lys | thr | ala | glu | ala | leu | 260 |
| gly | leu | ala | arg | arg | leu | tyr | gly | glu | gly | tyr | ala | pro | arg | ser | leu | val | ser | gly | leu | 280 |
| leu | glu | val | phe | arg | glu | glu | gly | leu | tyr | ala | phe | gly | leu | ala | gly | thr | pro | leu | pro | 300 |
| ala | pro | pro | gln | ala | leu | ile | ala | ala | met | thr | ala | leu | asp | glu | ala | met | glu | arg | leu | 320 |
| ala | arg | arg | ser | asp | ala | leu | ser | leu | glu | val | ala | leu | leu | glu | ala | gly | arg | ala | leu | 340 |
| ala | ala | glu | ala | leu | pro | gln | pro | thr | gly | ala | pro | ser | pro | glu | val | gly | pro | lys | pro | 360 |
| glu | ser | pro | pro | thr | pro | glu | pro | pro | arg | pro | glu | glu | ala | pro | asp | leu | arg | glu | arg | 380 |
| trp | arg | ala | phe | leu | glu | ala | leu | arg | pro | thr | leu | arg | ala | phe | val | arg | glu | ala | arg | 400 |
| pro | glu | val | arg | glu | gly | gln | leu | cys | leu | ala | phe | pro | glu | asp | lys | ala | phe | his | tyr | 420 |
| arg | lys | ala | ser | glu | gln | lys | val | arg | leu | leu | pro | leu | ala | gln | ala | his | phe | gly | val | 440 |
| glu | glu | val | val | leu | val | leu | glu | gly | glu | lys | lys | lys | ala |     |     |     |     |     |     | 454 |

FIG.4F

|         |   | ATP site |  |
|---------|---|----------|--|
| E.coli  | MSYQVLARKWRPOTFADVVGQEHVLTALANGLSLGRIHHAYLFSGTRGVGKTSIARLLAK  | 60       |  |
| H.inf.  | .....K.....II.....KDN.L.....F..                               | 60       |  |
| B.sub.  | ....A.Y.VF...R.E.....ITKT.Q.A.LQKKFS.....P.T....A.KIF...      | 60       |  |
| C.cres. | DA.T.....Y.R..E.LI...AMVRT...AF.T...A..FMLT.V.....TT.....R    | 113      |  |
| M.gen.  | -MH..FYQ.Y..IN.KQTL...SIRKI.V.AINRDKLPNG.I...E..T...TF.KII... | 59       |  |
| T.th.   | --VSA.Y.RF..L..QE.....KEP.LKAIRE..LAQ.....P.....TT.....M      | 58       |  |

|         | Zn <sup>++</sup> finger                                     |     |
|---------|---|-----|
|         | * * *   |     |
| E.coli  | GLNCET---GITATPCGVCDNCREIEQGRFVDLIEIDAASRTKVEDTRDLLDNVQYAPA | 116 |
| H.inf.  | ....VH----V.....E.E..KA...N.I.....E.....K.V                 | 116 |
| B.sub.  | AV...H---APVDE..NE.AA.KG.TN.SIS.V.....NNG.DEI..IR.K.KF..S   | 116 |
| C.cres. | A..Y..DTVK.PSVDLTTEGYH..S.IE..HM.VL.L.....DEM.E...G.R...V   | 173 |
| M.gen.  | AI..LN---WDQIDV.NS..V.KS.NTNSAI.IV.....KNGIN.I.E.VE..FNH.F  | 115 |
| T.th.   | AVG.QG-----EDP.....PH.QAVQR.AHP.VVD.....NNS...V.E.RERIHL..L | 112 |

|         |  |     |
|---------|--|-----|
| E.coli  | RGRFKVYLIDEVHMLSRHSFNALLKTLLEPPPEHVKFLLATDPQKLPVTILSRCLQFHLK | 176 |
| H.inf.  | V.....I.....IGA.....CI.I...E.H.I.L..I...QR.DE.               | 176 |
| B.sub.  | EA.Y...I.....TAA.....P.A..IF...EIR.V.....QR.D.R              | 233 |
| C.cres. | TFKK...IL..A...TTQ.WGG.....S.PY.L.IFT..EFN.I.L.....QS.FF..   | 175 |
| M.gen.  | SAPR..FIL..A....KSA.....P..L.VF...E.ERM.P.....TQH.RFR        | 172 |

FIG.5A

|         |   |     |
|---------|---|-----|
| E.coli  | ALDVEQIRHQLEHILNEEHIAHEPRALQLLARAAGEGSLRDALSLTDQAIASGDQ--VST  | 234 |
| H.inf.  | ...ET..SQH.A...TQ.N.PF.DP..VK..K..Q..I..S.....M..R.--.TN      | 234 |
| B.sub.  | RITSQA.VGRMNK.VDA.QLQV.EGS.EII.S..H.GM.....L....SFSGDI--LKV   | 234 |
| C.cres. | RVEPDVLVKHFDR.SAK.GARI.MD..A.I.....V..G...L....VQTERGQT.TS    | 293 |
| M.gen.  | KITSDL.LER.ND.AKK.K.KI.KD..IKI.DLSQ.....G...L..LAI.LIVKKL.LL  | 235 |
| T.th.   | R.TE.E.AFK.RR..EAVGREA.EE..L....L.D.A...E..LERFLLLEGP---LTR   | 229 |
| E.coli  | QAVSAMLGTLDDDDQALSIVEAMVEANGERVMA LINEAAARGIEWEALLVEMGLLHRIAM | 294 |
| H.inf.  | NV..N...L...NYSVDILY.LHQG...LL.RTLQRV.DAAGD.DK..G.CAEK..Q..L  | 294 |
| B.sub.  | EDALLIT.AVSQLYIGK.AKSLHDK.VSDALETL..LLQQ.KDPAK.IED.IFYFRDMLL  | 294 |
| C.cres. | TV.RD...LA.RS.TIA.Y.HVMAGKTKDALEGFRALWGF.ADPVVMLDV.DHC.AS.V   | 353 |
| M.gen.  | MLKKHLISLIEMQNL.L.KQFYQ.I                                     | 260 |
| T.th.   | KE.ERA..SPPGTGVAEIAASLARGKTAEALG.ARRLYGE.YAPRS.VSGL.EVFREGLY  | 289 |

FIG.5B

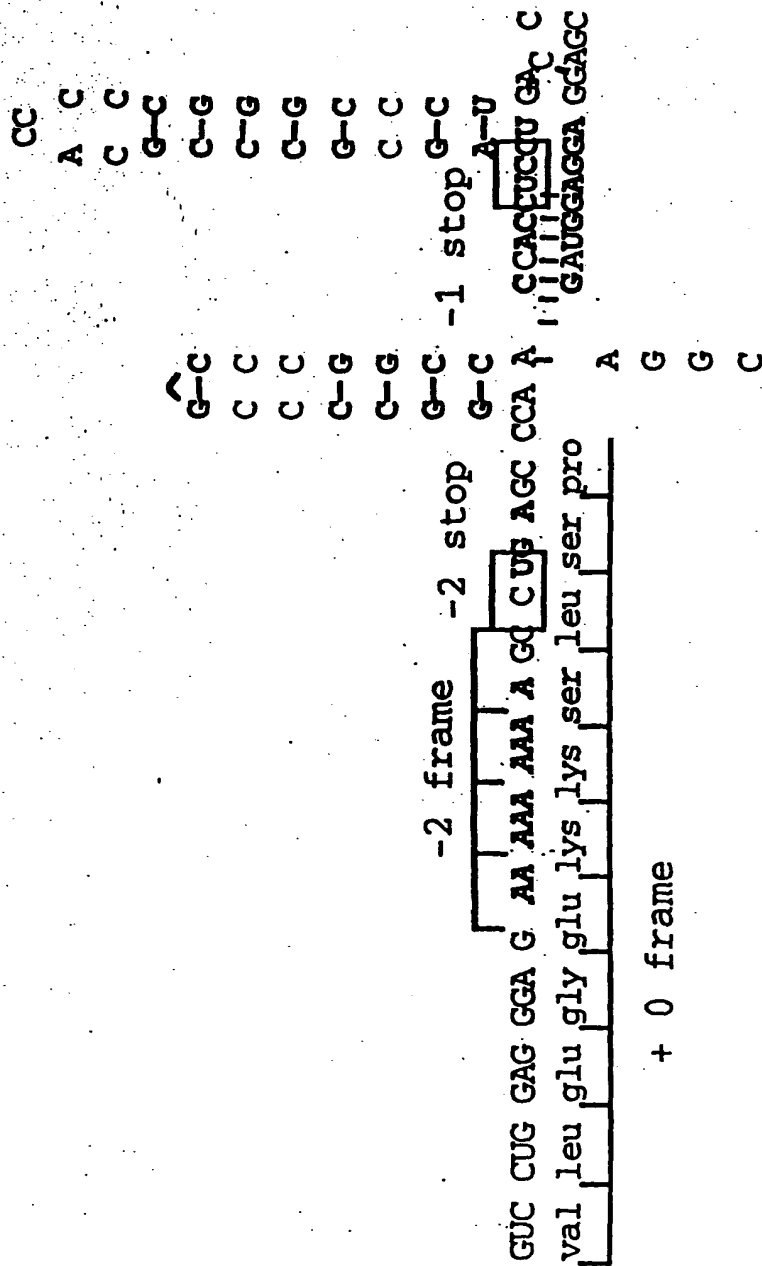


FIG.6

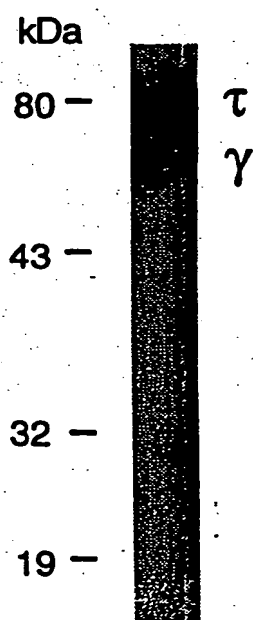
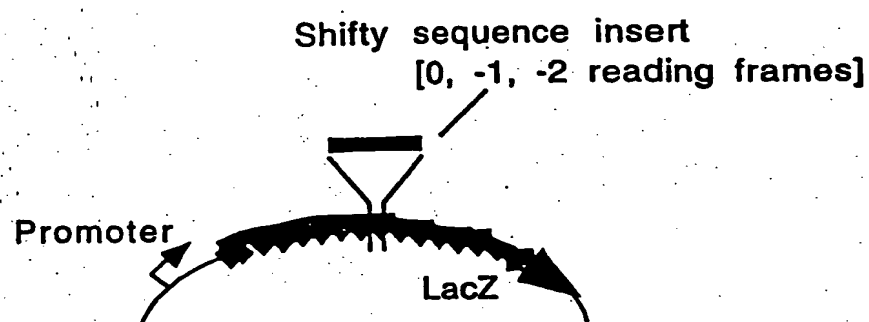


FIG. 7

# FIG.8A



|                 | Reading frame | Blue | White |
|-----------------|---------------|------|-------|
| Shifty sequence | 0             | +    |       |
|                 | - 1           | +    |       |
|                 | - 2           | +    |       |
| Mutant sequence | 0             | ++   |       |
|                 | - 1           |      | +     |
|                 | - 2           |      | +     |

# FIG.8B



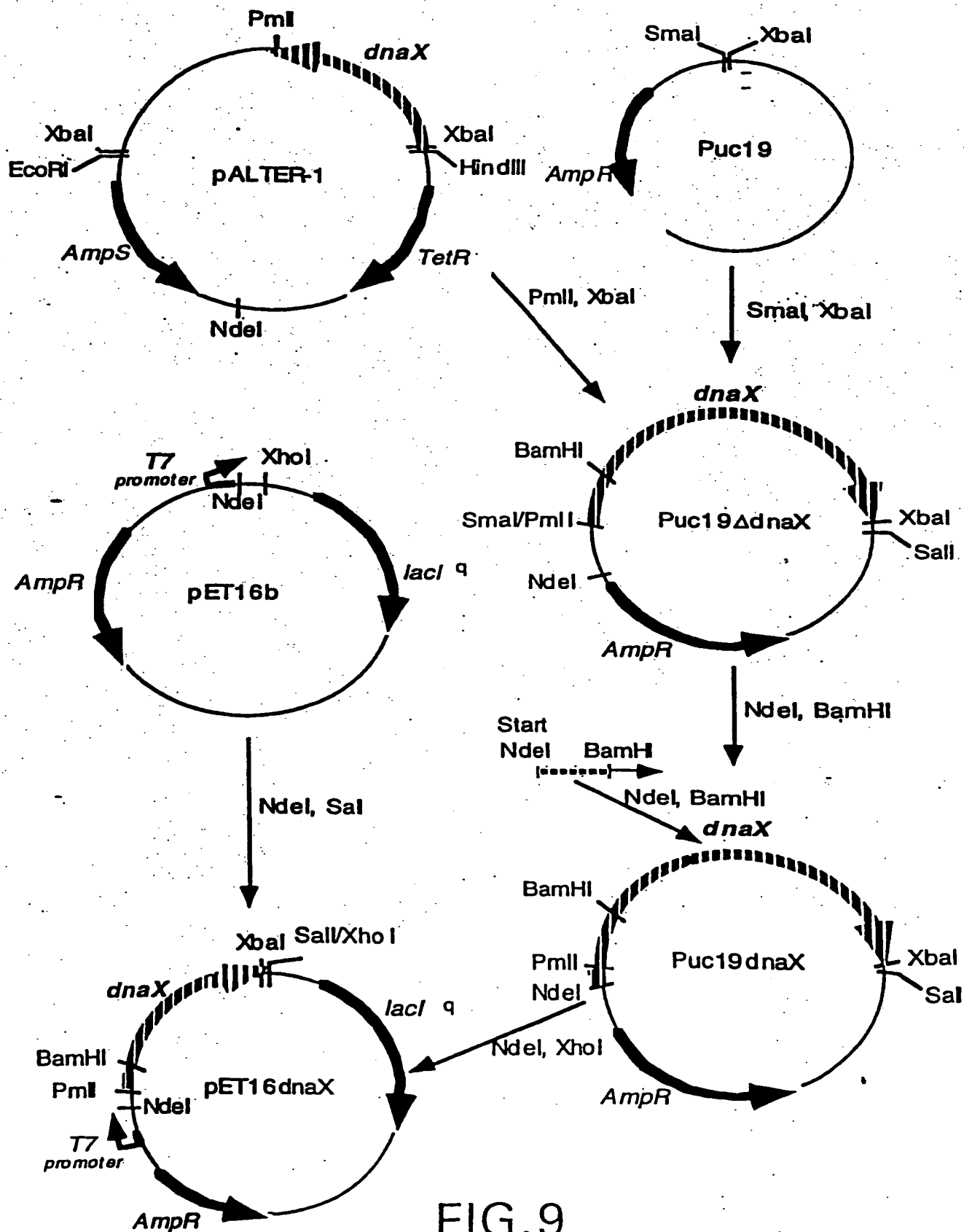


FIG.9

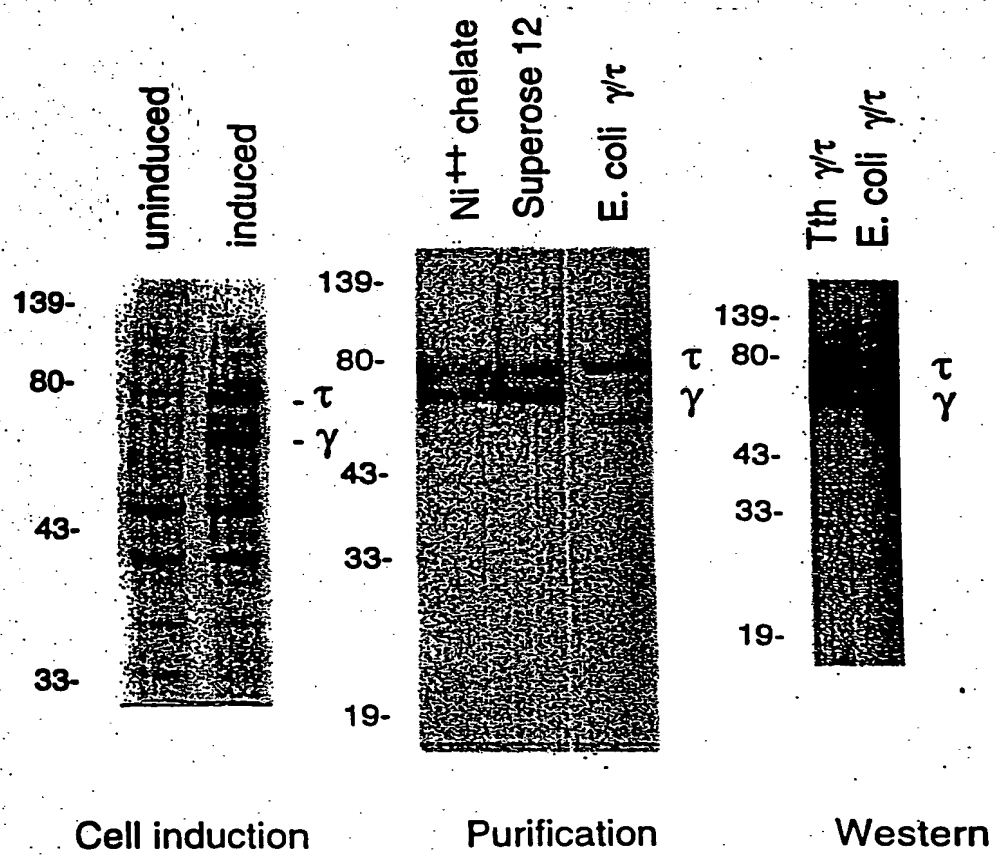


FIG.10A    FIG.10B    FIG.10C

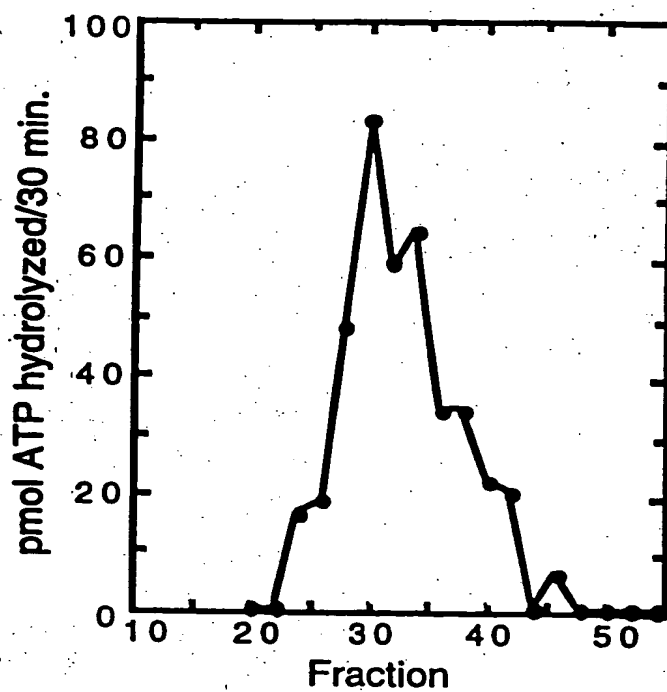


FIG. 11A

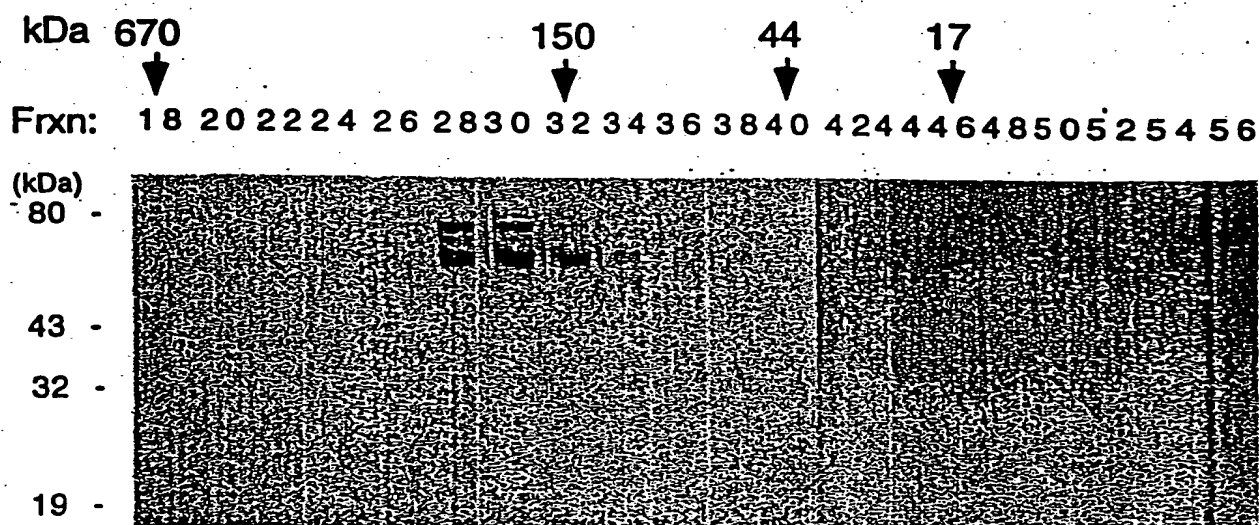


FIG. 11B

FIG.12A

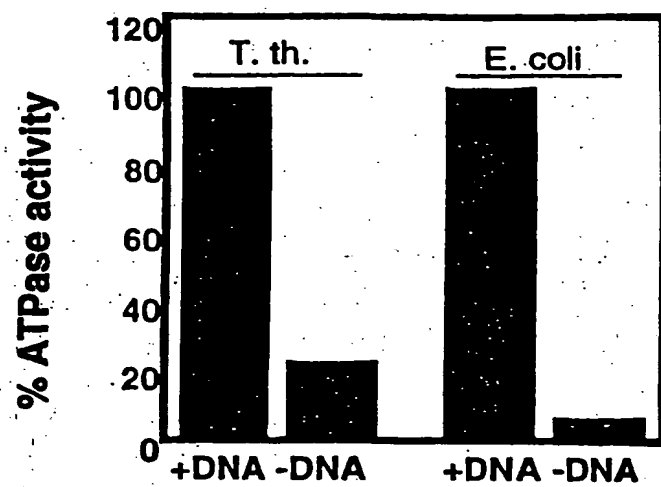


FIG.12B

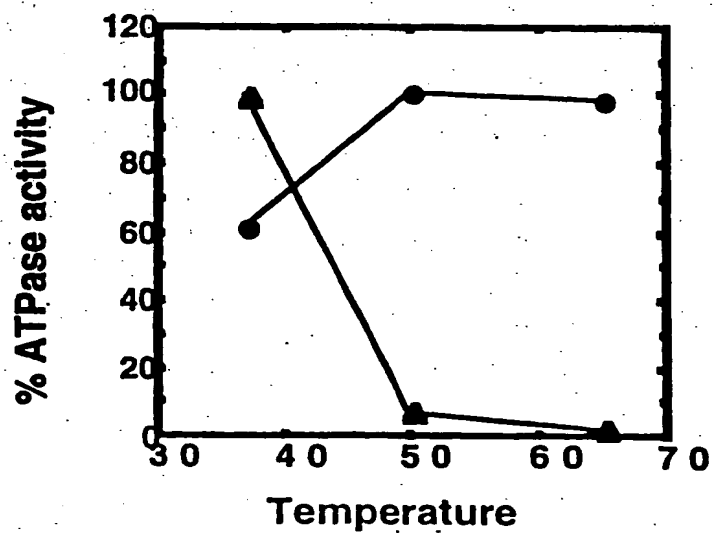
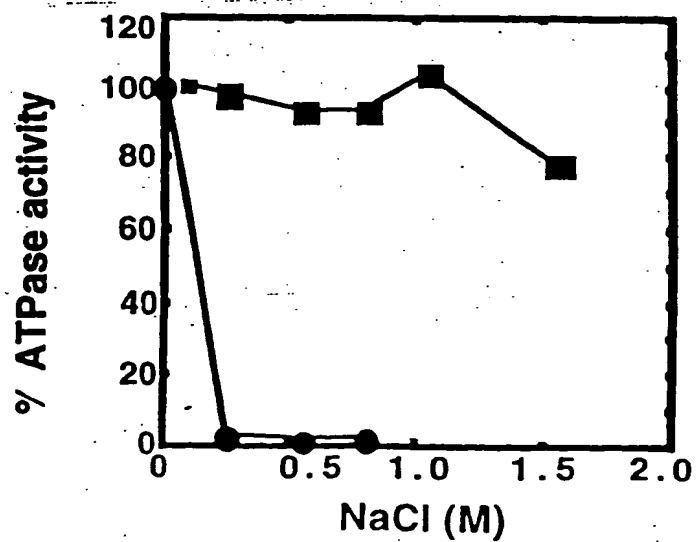
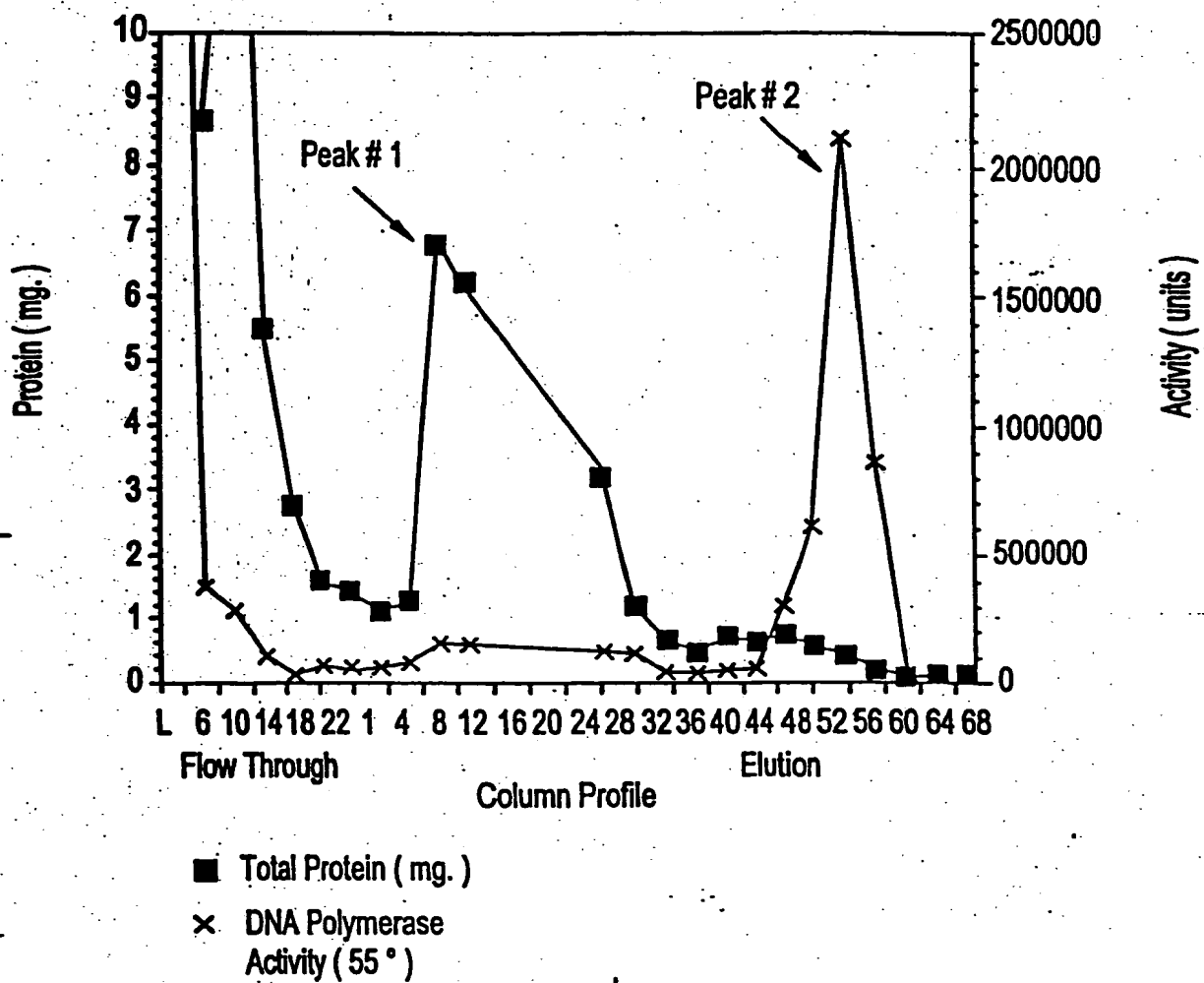


FIG.12C



# FIG.13A

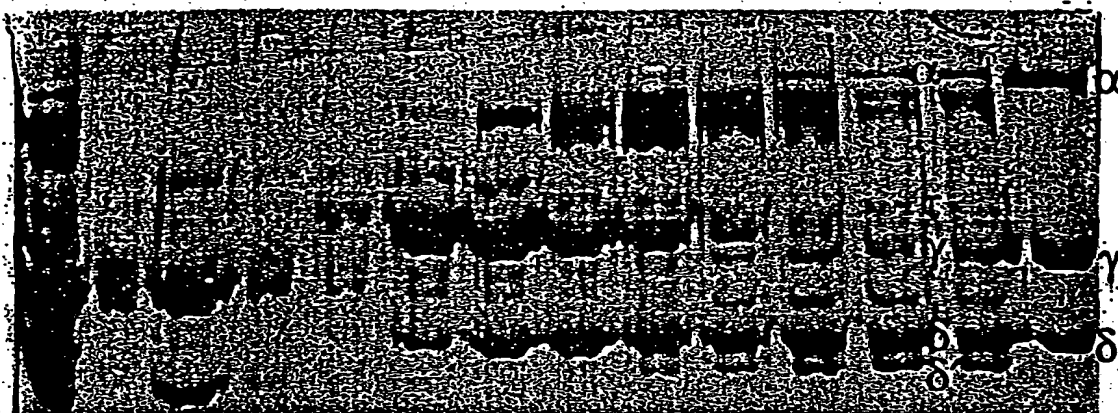


# FIG.13B

ATP Agarose Step Column

FIG.14A

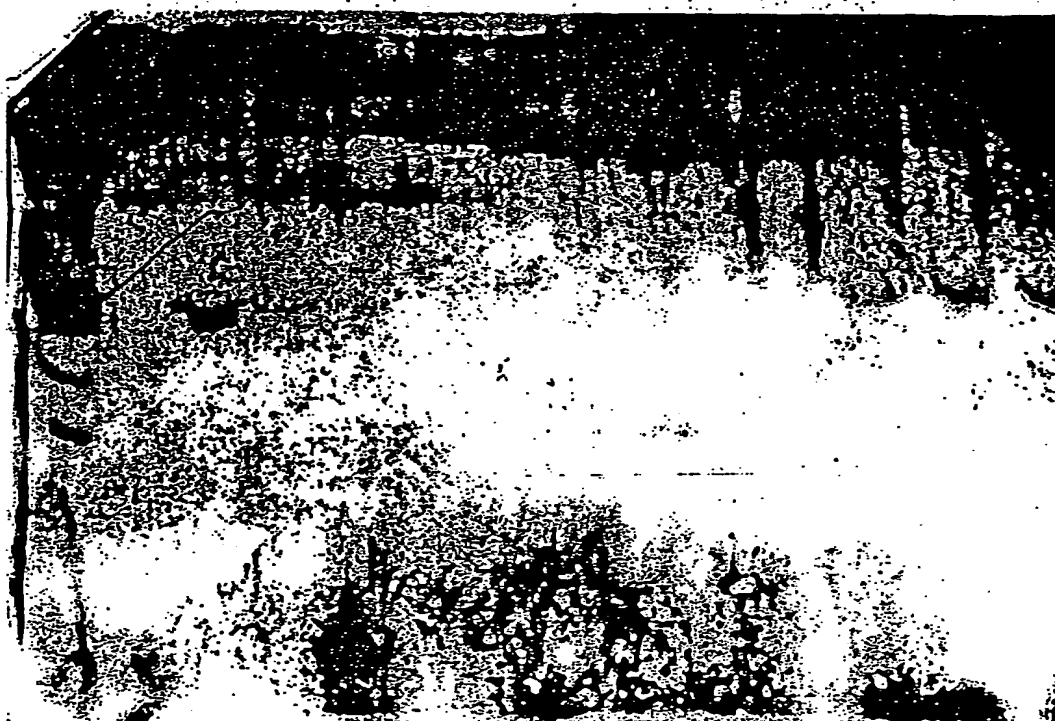
load FT 9 10 11 12 13 14 15 16 17 18 19 <sup>E. coli</sup>  
<sub>α γ δ</sub>



↑                      ↑  
T.th                      E. coli  
subunits                  subunits

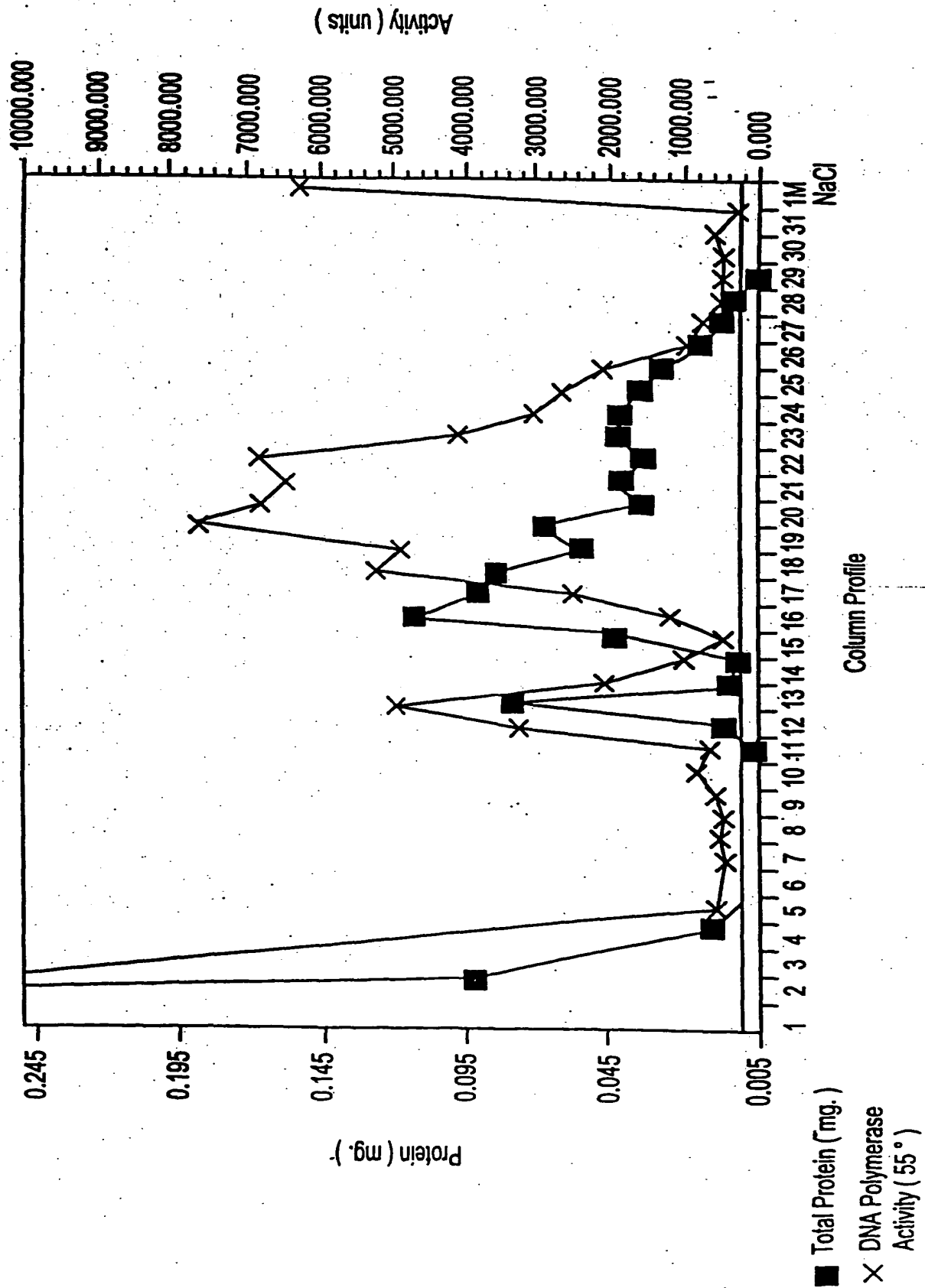
FIG.14B

load FT 9 10 11 12 13 14 15 16 17 18 19



← α

FIG.13C



Alignment of TTH1 with alphas subunits of other organisms.

|         |                                     |     |         |
|---------|-------------------------------------|-----|---------|
| E.coli. | DRYFLELIRTPDEESYLHAAVELAEARGLPVV    | 197 | (ID#72) |
| V.chol. | DHFYLELIRTPGRADEESYLHFDVAEQYDLPVV   | 197 | (ID#73) |
| H.inf.  | DHFYLAISRTPNEERYIQAAKLAERCPLV       | 197 | (ID#74) |
| R.prow. | DRFYFEIMRHDLPREEQFIENSYIQIASELSIPIV | 195 | (ID#75) |
| H.pyl.  | DDFYLEIMRHGILDQRFIDEQVIKMSLETGLKII  | 213 | (ID#76) |
| S.sp.   | DDYYLEIQDHGSVEDRLVNINLVKIAQELDIKIV  | 202 | (ID#77) |
| M.tub.  | DNYFLELMDHGLTIERRVRDGLLEIGRALNIPPL  | 220 | (ID#78) |
| T.th.   | FFIEIQNHGLSEQK                      |     | (ID#61) |

## FIG.15A

Alignment of TTH2 with alphas subunits of other organisms.

|         |   |     |         |
|---------|---|-----|---------|
| E.coli  | NKRRAKNGEPPLDIAAIPLDDKKSFDMLQRSETTAVFQLESRGMKD  | 618 | (ID#79) |
| V.chol. | NPRLKKAGKPPVRIEAIPLDDARSFRNLQDAKTTAVFQLESRGMKE  | 618 | (ID#80) |
| H.inf.  | NVRMVRGKPRVDIAAIPLDDPESFELLKRSETTAVFQLESRGMKD   | 618 | (ID#81) |
| R.prow. | CKKLLKEQGIKIDFDDMTFDDKITYQMLCKGKGVGFQFESIGMKD   | 624 | (ID#82) |
| H.pyl.  | LKIITQHKISVDFLSLDMDPPKVYKTIQSGDTVGFQIES-GMFQ    | 648 | (ID#83) |
| S.sp.   | QERKALQIRARTGSKKLPDDVKKTHKLLLEAGDLEGIFQLESQGMKQ | 643 | (ID#84) |
| M.tub.  | IDNVRANRGIDLDESVPDDKATYELLGRGDTLGVFQLDGGPMRD    | 646 | (ID#85) |
| T.th.   | RVELDYDALTLDD                                   |     | (ID#60) |

## FIG.15B



|  |      |
|--|------|
| ATGGGCGGGAGCTCCGCTTCGCCCACCTCCACCAGCACA  |      |
| CCCAGTTCTCCCTCCTGGACGGGGCGGCGAAGCTTTCCGA |      |
| CCTCCTCAAGTGGGTCAAGGAGACGACCCCCGAGGACCCC | 120  |
| GCCTTGGCCATGACCGACCACGGCAACCTCTTCGGGGCCG |      |
| TGGAGTTCTACAAGAAGGCCACCGAAATGGGCATCAAGCC |      |
| CATCCTGGGCTACGAGGCCTACGTGGCGGCGGAAAGCCGC | 240  |
| TTTGACCGCAAGCGGGGAAAGGGCCTAGACGGGGGCTACT |      |
| TTACCTCACCCTCCTCGCCAAGGACTTCACGGGGTACCA  |      |
| GAACCTGGTGCGCCTGGCGAGCCGGGCTTACCTGGAGGGG | 360  |
| TTTTACGAAAAGCCCCGGATTGACCGGGAGATCCTGCGCG |      |
| AGCACGCCGAGGGCCTCATCGCCCTCTCGGGGTGCCTCGG |      |
| GGCGGAGATCCCCCAGTTCATCCTCCAGGACCGTCTGGAC | 480  |
| CTGGCCGAGGCCCGGCTCAACGAGTACCTCTCCATCTTCA |      |
| AGGACCGCTTCTTCATCGAGATCCAGAACCACGGCCTCCC |      |
| CGAGCAGAAAAAGGTCAACGAGGTCCTCAAGGAGTTCGCC | 600  |
| CGAAAGTACGGCCTGGGGATGGTGGCCACCAACGACGGCC |      |
| ATTACGTGAGGAAGGAGGACGCCCGCGCCACGAGGTCTT  |      |
| CCTCGCCATCCAGTCCAAGAGCACCTGGACGACCCCGGG  | 720  |
| CGCTGGCGCTTCCCTGCGACGAGTTCTACGTGAAGACCC  |      |
| CCGAGGAGATGCGGGCCATGTTCCCGAGGAGGAGTGGGG  |      |
| GGACGAGCCCTTTGACAACACCGTGGAGATCGCCCGCATG | 840  |
| TGCAACGTGGAGCTGCCCATCGGGGACAAGATGGTCTACC |      |
| GAATCCCCCGCTTCCCCCTCCCCGAGGGGCGGACCGAGGC |      |
| CCAGTACCTCATGGAGCTCACCTTCAAGGGGCTCCTCCGC | 960  |
| CGCTACCCGGACCGGATCACCGAGGGCTTCTACCGGGAGG |      |
| TCTTCCGCCTTTTGGGGAAGCTTCCCCCCCACGGGGACGG |      |
| GGAGGCCTTGGCCGAGGCCTTGGCCCAGGTGGAGCGGGAG | 1080 |
| GCTTGGGAGAGGCTCATGAAGAGCCTCCCCCTTTGGCCG  |      |
| GGGTCAAGGAGTGGACGGCGGAGGCCATTTTCCACCGGGC |      |
| CCTTTACGAGCTTTCGGTGATAGAGCGCATGGGGTTTCCC | 1200 |
| GGCTACTTCTCATCGTCCAGGACTACATCAACTGGGCCC  |      |
| GGAGAAACGGCGTCTCCGTGGGGCCCGGCAGGGGGAGCGC |      |
| CGCCGGGAGCCTGGTGGCCTACGCCGTGGGGATCACCAAC | 1320 |
| ATTGACCCCTCCGCTTCGGCCTCCTCTTTGAGCGCTTCC  |      |
| TGAACCCGGAGAGGGTCTCCATGCCCGACATTGACACGGA |      |
| CTTCTCCGACCGGGAGCGGGACCGGGTGATCCAGTACGTG | 1440 |
| CGGGAGCGCTACGGCGAGGACAAGGTGGCCCAGATCGGCA |      |
| CCCTGGGAAGCCTCGCCTCCAAGGCCGCCCTCAAGGACGT |      |
| GGCCCCGGGTCTACGGCATCCCCACAAGAAGGCGGAGGAA | 1560 |
| TTGGCCAAGCTCATCCCGGTGCAGTTCGGGAAGCCCAAGC |      |
| CCCTGCAGGAGGCCATCCAGGTGGTGCCGGAGCTTAGGGC |      |
| GGAGATGGAGAAGGACCCCAAGGTGCGGGAGGTCTTCGAG | 1680 |
| GTGGCCATGCGCCTGGAGGGCCTGAACCGCCACGCCTCCG |      |
| TCCACGCCCGCGGGGTGGTGATCGCCGCCGAGCCCCTCAC |      |
| GGACCTCGTCCCCCTCATGCGCGACCAGGAAGGGCGGCCC | 1800 |
| GTCACCCAGTACGACATGGGGGCGGTGGAGGCCTTGGGGC |      |
| TTTTGAAGATGGACTTTTGGGCCTCCGCACCCTCACCTT  |      |

FIG. 16A

|  |      |
|--|------|
| CCTGGACGAGGTCAAGCGCATCGTCAAGGCGTCCCAGGGG | 1920 |
| GTGGAGCTGGACTACGATGCCCTCCCCCTGGACGACCCA  |      |
| AGACCTTCGCCCTCCTCTCCCGGGGGGAGACCAAGGGGGT |      |
| CTTCCAGCTGGAGTCGGGGGGGATGACCGCCACGCTCCGC | 2040 |
| GGCCTCAAGCCGCGGCGCTTTGAGGACCTGATCGCCATCC |      |
| TCTCCCTCTACCGCCCCGGGCCCATGGAGCACATCCCCAC |      |
| CTACATCCGCGCCACCACGGGCTGGAGCCCGTGAGCTAC  | 2160 |
| AGCGAGTTTCCCCACGCCGAGAAGTACCTAAAGCCCATCC |      |
| TGGACGAGACCTACGGCATCCCCGTCTACCAGGAGCAGAT |      |
| CATGCAGATCGCCTCGGCCGTGGCGGGGTACTCCCTGGGC | 2280 |
| GAGGCGGACCTCCTGCGGCGGTCCATGGGCAAGAAGAAGG |      |
| TGGAGGAGATGAAGTCCACCGGGAGCGCTTCGTCCAGGG  |      |
| GGCCAAGGAAAGGGGCGTGCCCGAGGAGGAGGCAACCGC  | 2400 |
| CTCTTTGACATGCTGGAGGCCTTCGCCAACTACGGCTTCA |      |
| ACAAATCCCACGCTGCCGCCTACAGCCTCCTCTCCTACCA |      |
| GACCGCCTACGTGAAGGCCCACTACCCCGTGAGTTTCATG | 2520 |
| GCCGCCCTCCTCTCCGTGGAGCGGCACGACTCCGACAAGG |      |
| TGGCCGAGTACATCCGCGACGCCCGGGCCATGGGCATAGA |      |
| GGTCCTTCCCCCGGACGTCAACCGCTCCGGGTTTGACTTC | 2640 |
| CTGGTCCAGGGCCGGCAGATCCTTTTCGGCCTCTCCGCGG |      |
| TGAAGAACGTGGGCGAGGCGGCGGCGGAGGCCATTCTCCG |      |
| GGAGCGGGAGCGGGGCGGCCCTACCGGAGCCTCGGCGAC  | 2760 |
| TTCTCAAGCGGCTGGACGAGAAGGTGCTCAACAAGCGGA  |      |
| CCCTGGAGTCCCTCATCAAGGCGGGCGCCCTGGACGGCTT |      |
| CGGGGAAAGGGCGCGGCTCCTCGCCTCCCTGGAAGGGCTC | 2880 |
| CTCAAGTGGGCGGCCGAGAACCGGGAGAAGGCCCGCTCGG |      |
| GCATGATGGGCCTCTTCAGCGAAGTGAGGAGCCGCCTTT  |      |
| GGCCGAGGCCGCCCCCTGGACGAGATCACCCGGCTCCGC  | 3000 |
| TACGAGAAGGAGGCCCTGGGGATCTACGTCTCCGGCCACC |      |
| CCATCTTGCGGTACCCCGGGCTCCGGGAGACGGCCACCTG |      |
| CACCCTGGAGGAGCTTCCCCACCTGGCCCGGGACCTGCCG | 3120 |
| CCCCGGTCTAGGGTCTCCTTGCCGGGATGGTGGAGGAGG  |      |
| TGGTGCAGCAAGCCCAAAAGAGCGGCGGGATGATGGCCCG |      |
| CTTCGTCTCTCCGACGAGACGGGGGCGCTTGAGGCGGTG  | 3240 |
| GCATTGCGCCGGGCCTACGACCAGGTCTCCCCGAGGCTCA |      |
| AGGAGGACACCCCGTGCTCGTCCCTCGCCGAGGTGGAGCG |      |
| GGAGGAGGGGGCGGTGCGGGTGCTGGCCAGGCCGTTTGG  | 3360 |
| ACCTACGAGGAGCTGGAGCAGGTCCCCCGGGCCCTCGAGG |      |
| TGGAGGTGGAGGCCTCCCTCCTGGACGACCGGGGGGTGGC |      |
| CCACCTGAAAAGCCTCCTGGACGAGCACGCGGGGACCCTC | 3480 |
| CCCCTGTACGTCCGGGTCCAGGGCGCCTTCGGCGAGGCC  |      |
| TCCTCGCCCTGAGGGAGGTGCGGGTGGGGGAGGAGGCTGT |      |
| AGGCGGCCGCGTGGTTCCGGGCCTACCTCCTGCCCGACCG | 3600 |
| GGAGGTCCTTCTCCAGGGCGGCCAGGCGGGGGAGGCCAG  |      |
| GAGGCGGTGCCCTTCTAGGGGGTGGGCCGTGAGACCTAGC |      |
| GCCATCGTTCTCGCCGGGGGCAAGGAGGCCTGGGCCCGAC | 3720 |
| CCCTTTTGG                                |      |

FIG. 16B

|   |      |
|---|------|
| MGRELRF AHLHQHTQFSLLDGAPKLSDLLKWVEETTPEDP   |      |
| ALAMTDHGNLFGAVEFYKKATEMGIKPILGYEAYVAAESR    |      |
| FDRKRGKGLDGGYFHLTLLAKDFTGYQNLVRLASRAYLEG    | 120  |
| FYEKPRIDREILREHAEGLIALSGCLGAEIQFILQDRLD     |      |
| LAEARLNEYLSIFKDRFFIEIQNHGLPEQKKVNEVLKEFA    |      |
| RKYGLGMVATNDGHYVRKEDARAHEVLLAIQSKSTLDDPG    | 240  |
| ALALPCEEFYVKTPEEMRAMFPEEEEVGGRSPLTTPWRSPPH  |      |
| VQRGAAIGTRWSTRIPRFPPLPEGRTEAQYLMELTFKGLLR   |      |
| RYPDRI TE GFYREVFR LSGKLPPHGDGEALAEALAQVERE | 360  |
| AWERLMKSLPPLAGVKEWTAEAI FHRALYELSAIERMGFP   |      |
| GLLPHRPGLHQLGPEKGVSVGPGRGGAAGSLVAYAVGITN    |      |
| IDPLRFGLL FERFLNPERVSMPDIDTDFSDRERDRVIQYV   | 480  |
| RERYGEDKVAQIGTLGSLASKAALKEVARVYGI PRKKAEE   |      |
| LAKLIPVQFGKPKPLQEAIQVVP ELRAEMEKDPKVREVLE   |      |
| VAMRLEGLNRHASVHAGRGGVFSEPLTDLVPLCATRKGGP    | 600  |
| YTQYDMGAVEALGLLKMDFLGLRTLTLFLDEVKRIVKASQG   |      |
| VELDYDALPLDDPKTFALLSRGETKGVFQLESGGMTATLR    |      |
| GLKPRRFEDLIAILSLYRPGPMEHIPTYIRRHGLEPVSY     | 720  |
| SEFPHAEKYLKPILDETYGIPVYQEQIMQIASAVAGYSLG    |      |
| EADLLRRSMGKKKVEEMKSHRERFVQGAKEGVP EEEANR    |      |
| LFDMLEAFANYGFNKSHAAAYSLLSYQTAYVKAHYPVEFM    | 840  |
| AALLSVERHDSKVAEYIRDARAMGIEVLPPDVNRSGFDF     |      |
| LVQGRQILFGLSAVKNVGEAAAEAILRERERGGPYRSLGD    |      |
| FLKRLDEKVLNKRTLES LIKAGALDGFGERARLLASLEGL   | 960  |
| LKWAAENREKARSGMMGLFSEVEEPPLAEAAPLDEITRLR    |      |
| YEKEALGIYVSGHPILRYPGLRETATCTLEELPHLARDLP    |      |
| PRSRVLLAGMV EEVVRKPTKSGGMMARFVLSDETGALEAV   | 1080 |
| AFGRAYDQVSPRLKEDTPVLVLAEVEREEGGVRVLAQAVW    |      |
| TYQELEQVPRALEVEVEASLPDDRGV AHLKSLLD E HAGTL |      |
| PLYVRVQGA FG EALLALREVRVGE EALGALEAAGFPAYLL | 1200 |
| PNREVS PRLTGSGGPRGRALSTGLALKTYP IALPGGNEAL  |      |
| ARPLL                                       |      |

FIG. 16C

|          | Start1   | Start2   | 3'-Exo I    |
|----------|--|--|-------------|
| T.th.    | VERVVRTLLDGRFLLEEGVGLWERYPPFLEGEAVVLDLETTGLAG-----LDEVIEUGLLRLEGG---RRLPF          |  |             |
| D.rad.   |  | PWPQDVVVFDDLETTGFSPA-----SAAIVEIGAVRIVGGQIDETLKF   |             |
| Bac.sub. | HGIKMIYMEANLVDDGVPIAYNAAHRLLEEETVVFDDVETTGLSAV-----YDTIIELAAVKVKGGE--IIDKF         |  |             |
| H.inf.   |  | MINPNRQIVLDTTETTMNQLGAHYEGHCIIIEIGAVELINRR-YTGNNX  |             |
| E.c.     |  | MSTAITRQIVLDTTETTMNQLGAHSEGHKIIIEIGAVEVNNRR-LTGNNF |             |
| H.pyl.   | NLEYLKACGLNFIETSENLTILKNLKTPLKDEVFSFIDLETTGSCPI-----KHEILEIGAVQVKGE--IINRF         |  |             |
|          |  |  | 3'-Exo II   |
| T.th.    | QSLVR-PLPP---AEARSWNLT---GIPREALLEAPSLEEVLEKAYPLRGDATLVIHNAAFDLGFL-RPALEGLG        |  |             |
| D.rad.   | ETLVR-PTRPDGSMLSI PWQAQRVHGISDEMVRRA PAKDVL P DFFDFV DGS AVVAHNVSFDGGFM-RAGAERLG   |  |             |
| Bac.sub. | EAFAN-PHRP---LSATIIELT---GITDDMLQDAPD VDVIRDFREWIGDDILVAHNASFDMGFL-NVAYKKLL        |  |             |
| H.inf.   | HIYIK-PDRP---XDPDAIKVH---GITDEMLADKPEFKEVAQDFLDYINGAELLIHNAFFDVGFM-DYEFRRKLN       |  |             |
| E.c.     | HVYLK-DRLV---DPEAFGVH---GIAVDFLDKPTFAEVAVEFMDYIRGAELV IHNAAFD I GFM-DYEFSLLK       |  |             |
| H.pyl.   | ETLVKVKSV P-----DYIAELT---GITYEDTLNAPSAHEALQELRLFLGNSVFVAHNANFDYNFLGRYFVEKLH       |  |             |
|          |  |  | 3'-Exo IIIC |
| T.th.    | -----YRLENPVVDSLRLARRGLPGLRRYGLDALSEVLELPRRT--CHRALEDVERTLAVVHEVYVYMLT-----SG      |  |             |
| D.rad.   | -----LSWAPERELCTMQLSRRAFP RERTHNLT VLAERLGL EFA PGGRHRSYGDVQVTAQAYLR LLELLG-----ER |  |             |
| Bac.sub. | E---VEKAKNPVIDTLELGRFLYPEFKNHRNLNTLCKKFDIELTQ--HHRAIYDTEATAVYLLKMLKDA A-----EK     |  |             |
| H.inf.   | -LNVKTDDICLVDTLQMAROMYPGKRN-NLDALCDRLGIDNSKRTLHGALLDAEILADVYLLMMTGGQTNLFDEEE       |  |             |
| E.c.     | RDIAKTNTFCKVTDLSLAVARKMFP GKR N-SLDALCARYEIDNSKRTLHGALLDAQI LAEVYLA MTGGQTSMAFAME  |  |             |
| H.pyl.   | -----CPLNLKLCTLDLSKRAILSMRY-SLSFLKELLGFGIEV--SHRAYADALASYKLFEICLLNLP--SYIKT        |  |             |

FIG.17

## FIG.18A

ATGGTGGAGCGGGTGGTGCGGACCCTTCTGGACGGGAGGT 40  
TCCTCCTGGAGGAGGGGGTGGGGCTTTGGGAGTGGCGCTA  
CCCCTTTCCCCTGGAGGGGGAGGCGGTGGTGGTCTTGAC 120  
CTGGAGACCACGGGGCTTGCCGGCCTGGACGAGGTGATTG  
AGGTGGGCCTCCTCCGCCTGGAGGGGGGAGGCGCCTCCC 200  
CTTCCAGAGCCTCGTCCGGCCCCCTCCCGCCCGCCGAAGCC  
CGTTCGTGGAACCTCACCGGCATCCCCGGGAGGCCCTGG 280  
AGGAGGCCCCCTCCCTGGAGGAGGTTCTGGAGAAGGCCTA  
CCCCCTCCGCGGCGACGCCACCTTGGTGATCCACAACGCC 360  
GCCTTTGACCTGGGCTTCCTCCGCCCCGGCCTTGGAGGGCC  
TGGGCTACCGCCTGGAAAACCCCGTGGTGGACTCCCTGCG 440  
CTTGGCCAGACGGGGCTTACCAGGCCTTAGGCGCTACGGC  
CTGGACGCCCTCTCCGAGGTCTTGGAGCTTCCCCGAAGGA 520  
CCTGCCACCGGGCCCTCGAGGACGTGGAGCGCACCCCTCGC  
CGTGGTGCACGAGGTATACTATATGCTTACGTCCGGCCGT 600  
CCCCGCACGCTTTGGGAACCTCGGGAGGTAG

MVERVVRTLLEDGRFLLEEGVGLWEWRYPFPLEGEAVVLD 40  
LETTGLAGLDEVIEVGLLRLEGGRRLPFQSLVRPLPPAEA  
RSWNLTGIPREALEEAPSLEEVLEKAYPLRGDATALVIHNA 120  
AFDLGFLRPALEGLGYRLNPVVDLRLARRGLPGLRRYG  
LDALSEVLELPRRTCHRALEDVERTLAVVHEVYMLTSGR 200  
PRTLWELGRZ

## FIG.18B

# Alignment of dnaA genes.

|          |                                   |                                 |                                   |     |
|----------|-----------------------------------|---------------------------------|-----------------------------------|-----|
| P. mar.  | MLEASWEK VQSSL--KQNLK--           | -----PSYE TWIRPTESG--FKN        | GELTLIAPNSFSSAW LKNYSQTIQETAE--   | 65  |
| Syn. sp. | MVSCENLWQQ ALAIL--ATQLTK--        | -----PAFD TWIKASVLIS--LGD       | GVATIQVENGFVLNH LQSYGPLLMEVLT--   | 67  |
| B. sub.  | MENILDWNQ ALAQI--EKQLSK--         | -----PSFE TWKSTKAHS--LQG        | DTLTTAPNEFARDW LESRYLHLIADTTY--   | 67  |
| M. tub.  | MTDDPGSGFTTVWNA WSELNGDPKVDGDP    | SSDANLSAPLTPQQR AMLNLVQPLT--IVE | GFALLSVPSSEFVQNE IERHLRAPITDALS-- | 87  |
| T. th.   | MSHEAVWQH VLEHI--RRSITE--         | -----VEFH TWFERIRPIG--IRD       | GVLELAVPTSEALDW IRRHYAGLIQEGPR--  | 66  |
| E. coli  | MSLSLWQQ CLARL--QDELPA--          | -----TEFS MWIRPLQAE--LSD        | NTLALYAFNRFVLDW VRDKYLANNINGLLT-- | 64  |
| T. mar.  | MKER ILQEI--KTRVNR--              | -----KSWE LMFSSFDVKS--IEG       | NKVVFSVGNLFIKEW LEKKYYSVLKAVK--   | 61  |
| H. pyl.  | MDTNNNIEKE ILALVKQNPVSL--         | -----IEVE NYFSQIKYNPNASKS       | DLAFFYAPNQVLCIT ITAKYGALLKEILSQ   | 72  |
| P. mar.  | EIFG---EPVTVHVK VKANAESSDEHYSSA P | -----ITPPLEASPGSV DSSGSSLRSK    | -----KTLPLNLRYVFN                 | 130 |
| Syn. sp. | DLTG---QEITVKLI TDGLEPHS---LIGQ E | -----SSLPMETTP                  | -----KNATALNGKYTF                 | 115 |
| B. sub.  | ELTG---EELSIFV IPQNUJVEDFMKPKQ    | VKKAVKEDTSDFPQN                 | -----MLNPKYTFDT                   | 119 |
| M. tub.  | RRLGH-QIQLGVRIA PPATDEADDTTVPPS   | ENPATTSPDTTDND                  | EIDDSAAARGDNQHS WPSYFTIERPHNTDSA  | 176 |
| T. th.   | LLGAQ-APRFELRVV PGVVQEDIFQPPPS    | PPAQAP--                        | -----EDTFKT                       | 108 |
| E. coli  | SFCGADAPQLRFEVG TKPVTQTTPQAAVTSN  | VAAPQVAQTQPORA                  | APSTRSGWDNVPAPA EP                | 140 |
| T. mar.  | VVLG---NDATFEIT YEAFEPHSSYSEPLV   | KRAVLLTP                        | -----LNPDYTFEN                    | 106 |
| H. pyl.  | NKVG-MHLAHSVDVR IEVAPKIQINAQSN    | NYKAIKTS                        | -----VKDSYTFEN                    | 118 |
| P. mar.  | FVVGPNRMAHAAAM AVAESPGREENPLFI    | CGGVGLGKTHLMQAI                 | GHYRLEIDPGAKVSY VSTETFTNDLIL--A   | 217 |
| Syn. sp. | FVVGPTNRMHAASL AVAESPGREENPLFL    | CGGVGLGKTHLMQAI                 | AHYRLEMYPNAKVY VSTERFTNDLIT--A    | 202 |
| B. sub.  | FVIGSGNRFHAASL AVAEPAPAKAVNPLFI   | YGGVGLGKTHLMHAI                 | GHYVIDHNPSAKVY LSSEKFTNEFIN--S    | 206 |
| M. tub.  | FVIGASNRFHAHAAL ALAEPAPARAVNPLFI  | WGESGLGKTHLLHAA                 | GNYAQRLFGMRVKY VSTEEFTNDFIN--S    | 263 |
| T. th.   | SWMGPTTPWPHGGAV AVAESPGRAVNPLFI   | YGGRGLGKTYLMHAV                 | GPLRAKRFPHMRLEY VSTETFTNELINRPS   | 196 |
| E. coli  | FVEGKSNQLARAAAR QVADNPGGAVNPLFL   | YGGTGLGKTHLLHAV                 | GNGIMARKPNKVVY MHSERFVQDMVK--A    | 227 |
| T. mar.  | FVVGPGNSFAYHAAL EVAKHPGR--YNPLFI  | YGGVGLGKTHLLQSI                 | GNYVQNEPDLRVMY ITSEKFLNDLVD--S    | 193 |
| H. pyl.  | FVVGSCNNTVYEIAK KVAQSDTPPPVNPVLF  | YGGTGLGKTHILNAI                 | GNHALEK--HKKVVL VTSEDFLTDFLK--H   | 203 |

FIG. 19A

|         |                 |                 |                 |                 |                  |                  |     |
|---------|-----------------|-----------------|-----------------|-----------------|------------------|------------------|-----|
| P.mar.  | AADLLVDDIQFIEG  | KEYTQEEFFHTFNAL | HDAGSQIVLASDRPP | SQIPRLQERLMSRFS | MGLIADVQAPDLETR  | MAILQKKAHERVGL   | 307 |
| Syn.sp. | SADFLIDDIOFIKG  | KEYTQEEFFHTFNAL | HEAGQVWVASDRAP  | QRIFGLQDRLISRFS | MGLIADIQVPDLETR  | MAILQKKAHYDRIRL  | 292 |
| B.sut.  | NVDVLLIDDIOFLAG | KEQTQEEFFHTFNAL | HEESQIVISSDRPP  | KEIPTLEDRLRSRFE | WGLITDITPDPDLETR | IAILRKKAKAEGLDI  | 296 |
| M.tub.  | DVDVLLVDDIQFIEG | KEGTQEEFFHTFNAL | HNANKQIVISSDRPP | KQATLEDRLRTRFE  | WGLITDVQPPDLETR  | IAILRKKQAQMERLAV | 353 |
| T.th.   | SVDDLVDVDDQFIAG | KERTQEEFFHTFNAL | YEAHKQIILSSDRPP | KDILTEARLSRFE   | WGLITNPPDLETR    | IAILKMNAS-SGPED  | 285 |
| E.coli  | SVDDLIDDIOFFAN  | KERSQEEFFHTFNAL | LEGNQOIILTSDRYP | KEINGVEDRLKSRFG | WGLITVAIEPPELETR | VAIIMKKADENDIRL  | 317 |
| T.mar.  | KVDILLIDDVQFLIG | KTGVQTELFHTFNEL | HDGKQIVICSOREP  | QKLSEFQDRLVSRFQ | MGLVAKLEPPDETR   | KSIARKMLEIEHGEL  | 283 |
| H.pyl.  | HCDFLLDDAQFLOQ  | KPKLEEEFFHTFNEL | HANSKQIVLISDRSP | KNIAGLEDRLKSRFE | WGITAKVMPDLETK   | LSIVKQKCOLNQITL  | 293 |

|         |                 |                 |                  |                 |                  |                 |     |
|---------|-----------------|-----------------|------------------|-----------------|------------------|-----------------|-----|
| P.mar.  | PRDLIOFIAGRFTSN | IRELEGALTRAIAFA | SITGLPMTVDISIAPM | LD----PNGQGVET  | PQVLDKVAEVFKVT   | PDEMRASRRR-PVS  | 392 |
| Syn.sp. | PKEVIEYIASHYTSN | IRELEGALIRAIAYT | SLSNVAMTVENIAPV  | LN----PPVEKVAAA | PETIITITVAQHYQLK | VEELLSNSRRR-EVS | 377 |
| B.sut.  | PNEVMLYIANQIDSN | IRELEGALIRVWAYS | SLINKDINADLAAEA  | LKOII-PSSKPKVIT | IKEIQRWVGQQFNIK  | LEDFKAKKRTK-SVA | 384 |
| M.tub.  | PDDVLELIASSIERN | IRELEGALIRVTAF  | SLNKTPIDKALAEIV  | LRDLI-ADANTMQIS | AATIMAATAEYFDIT  | VEELRGPGKTR-ALA | 441 |
| T.th.   | PEDALEYIARQVTSN | IREWEGALMRASPPA | SINGVELTRAVAACA  | LRHLR-P--RELEAD | PLEIIRKAAGPVRPE  | TPGGAHGERRKCEV  | 372 |
| E.coli  | PGEVAFFIAKRLRSN | VRELEGALNRVIANA | NFTGRAITIDFVREA  | LRDLL-A-LQEKLV  | IDNIQKTVAEYKIK   | VADLLSKRRSR-SVA | 404 |
| T.mar.  | PEEVINFVAENVDDN | LRRLRGAIKLLVYK  | ETTQKEVDLKEAAIL  | LKOFIKPNRVKAMD  | IDELTEIVAKVTGVP  | REEILSNSRNV-KAL | 372 |
| H.pyl.  | PEEVMEYIAQHISDN | IRQMEGAIKISVNA  | NLMNASIDINLAKTV  | LEDL--OKDHAEGSS | LENILLAVAQSLNLK  | SSEIKVSSRQK-NVA | 380 |

|         |                 |                 |                  |                 |                  |          |     |
|---------|-----------------|-----------------|------------------|-----------------|------------------|----------|-----|
| P.mar.  | QARQVGYLMRQGTN  | LSLPRIGDTFGGKDH | TTVMYAIEQVEKKLS  | S-----DPQIA     | SQVQKIRDLLQIDSR  | RKR----  | 461 |
| Syn.sp. | LARQVGYLMRQHTD  | LSLPRIGEAFGGKDH | TTVMYSCDKITQLQ   | K-----DWETS     | QTLTSLSHRINIAGQ  | APES---- | 447 |
| B.sut.  | FPRQIAMYLSREMTD | SSLPKIGEEFGGRDH | TTVIHAHEKISKLLA  | D-----DEQLQ     | QHVKEIKBQLK----  | -----    | 446 |
| M.tub.  | QSRQIAMYLCRELTD | LSLPKIGQAFG-RDH | TTVMYAQRKILSEMA  | E-----RREVF     | DHVKELTTRIRQSK   | R-----   | 507 |
| T.th.   | LPRQIAMYLVRELTP | ASLPEIGQLFGGRDH | TTVRYAIQKVQELAG  | KP-----DREVQ    | GLLRTLREACTDPVD  | NLWITCG  | 446 |
| E.coli  | RPRQWAMALAKELTN | HSLPEIGDAFGGRDH | TTVLHACRKIEQLRE  | E-----SHDIK     | EDFSNLIRTLSS---- | -----    | 467 |
| T.mar.  | TARRIGMYVAKNYLK | SSLRTIAEKEN-RSH | PVVVDSVKVKVDLSLL | KG-----NKQLK    | ALIDEVIGEISRRAL  | SG-----  | 440 |
| H.pyl.  | LARKLVVYFARLYTP | NFTLSLAQFLDLKDH | SSISKMYSGVKKMLE  | EEKSPFVLSLREEIK | NRLNELNDKKTAFN   | SE-----  | 457 |

FIG.19B

GTGTCGCACGAGGCCGTCTGGCAACACGTTCTGGAGCAĒA  
 TCCGCCGCAGCATCACCGAGGTGGAGTTCCACACCTGGTT  
 TGAAAGGATCCGCCCCCTTGGGGATCCGGGACGGGGTGCTG 120  
 GAGCTCGCCGTGCCACCTCCTTTGCCCTGGACTGGATCC  
 GCGGCCACTACGCCGGCCTCATCCAGGAGGGCCCTCGGCT  
 CCTCGGGGGCCAGGCGCCCCGGTTTGAGCTCCGGGTGGTG 240  
 CCGGGGGTTCGTAGTCCAGGAGGACATCTTCCAGCCCCCGC  
 CGAGCCCCCGGCCAAGCTCAACCCGAAGATACCTTTAA  
 AACTTCGTGGTGGGGCCCAACAACCTCCATGGCCCCACGGC 360  
 GCGCCGTGGCCGTGGCCGAGTCCCCCGGCCGGGCTACA  
 ACCCCCTCTTCATCTACGGGGGCCGTGGCCTGGGAAAGAC  
 CTACCTGATGCACGCCGTGGGCCCACTCCGTGCGAAGCGC 480  
 TTCCCCCACATGAGATTAGAGTACGTTTCCACGGAAACTT  
 TCACCAACGAGCTCATCAACCGGCCATCCGCGAGGGACCG  
 - GATGACGGAGTTCCGGGAGCGGTACCGCTCCGTGGACCTC 600  
 CTGCTGGTGGACGACGTCCAGTTCATCGCCGGAAAGGAGC  
 GCACCCAGGAGGAGTTTTTCCACACCTTCAACGCCCTTTA  
 CGAGGCCCAAGCAGATCATCCTCTCCTCCGACCGGCCG 720  
 CCAAGGACATCCTCACCCCTGGAGGCGCGCCTGCGGAGCC  
 GCTTTGAGTGGGGCCTGATCACCGACAATCCAGCCCCCGA  
 CCTGGAAACCCGGATCGCCATCCTGAAGATGAACGCCAGC 840  
 AGCGGGCCTGAGGATCCCGAGGACGCCCTGGAGTACATCG  
 CCCGGCAGGTCACCTCCAACATCCGGGAGTGGGAAGGGGC  
 CCTCATGCGGGCATCGCCTTTCGCCTCCCTCAACGGCGTT 960  
 GAGCTGACCCGCGCCGTGGCGGCCAAGGCTCTCCGACATC  
 TTCGCCCCAGGGAGCTGGAGGCGGACCCCTTGGAGATCAT  
 CCGCAAAGCGGCGGGACCAGTTCGGCCTGAAACCCCGGGA 1080  
 GGAGCTCACGGGGAGCGCCGCAAGAAGGAGGTGGTCCTCC  
 CCCGGCAGCTCGCCATGTACCTGGTGCGGGAGCTCACCCC  
 GGCCTCCCTGCCCCGAGATCGACCAGCTCAACGACGACCGG 1200  
 GACCACACCACGGTCCTCTACGCCATCCAGAAGGTCCAGG  
 AGCTCGCGGAAAGCGACCGGGAGGTGCAGGGCCTCCTCCG  
 CACCCTCCGGGAGGCGTGACATGA

FIG.20A



VSHEAVWQHVLHRRSITEVEFHTWFERIRPLGIRDGVL  
ELAVPTSFALDWIRRHAGLIQEGPRLPGAQAPRFELRVV  
PGVVVQEDIFQPPPSPPAQAPEDTFKTSWWGPTTPWPHG 120  
GAVAVAESPGRAYNPLFIYGGRLGKTYLMHAVGPLRAKR  
FPHMRLEYVSTETFTNELINRPSARDRMTEFRERYRSVDL  
LLVDDVQFIAGKERTQEEFFHTFNALYEAHKQIILSSDRP 240  
PKDILTLEARLRSRFEWGLITDNPAPDLETRIAILKMNAS  
SGPEDPEDALEYIARQVTSNIREWEGALMRASPFASLNGV  
ELTRAVAAKALRHRLPRELEADPLEIIRKAAGPVRPETPG 360  
GAHGERRKKEVVLPRQLAMYLVRCLTPASLPEIDQLNDDR  
DHTTVLYAIQKVQELAESDREVQGLLRTLREACT

FIG.20B

ATGAACATAACGGTTCCCAAAAACTCCTCTCGGACCAGC 40  
 TTTCCCTCCTGGAGCGCATCGTCCCCTCTAGAAGCGCCAA  
 CCCCCTCTACACCTACCTGGGGCTTTACGCCGAGGAAGGG 120  
 GCCTTGATCCTCTTCGGGACCAACGGGGAGGTGGACCTCG  
 AGGTCCGCCTCCCCGCCGAGGCCCAAAGCCTTCCCCGGGT 200  
 GCTCGTCCCCGCCAGCCCTTCTTCCAGCTGGTGCGGAGC  
 CTTCTGGGGACCTCGTGGCCCTCGGCCTCGCCTCGGAGC 280  
 CGGGCCAGGGGGGGCAGCTGGAGCTCTCCTCCGGGCGTTT  
 CCGCACCCGGCTCAGCCTGGCCCTGCCGAGGGCTACCCC 360  
 GAGCTTCTGGTGCCCGAGGGGGAGGACAAGGGGGCCTTCC  
 CCTCCGGACGCGGATGCCCTCCGGGGAGCTCGTCAAGGC 440  
 CTTGACCCACGTGCGCTACGCCGCGAGCAACGAGGAGTAC  
 CGGGCCATCTTCCGCGGGGTGCAGCTGGAGTTCTCCCCC 520  
 AGGGCTTCCGGGCGGTGGCCTCCGACGGGTACCGCCTCGE  
 CCTCTACGACCTGCCCCTGCCCCAAGGGTTCCAGGCCAAG 600  
 GCCGTGGTCCCCGCCCGAGCGTGGACGAGATGGTGCGGG  
 TCCTGAAGGGGGCGGACGGGGCCGAGGCCGTCTCGCCCT 680  
 GGGCGAGGGGGTGTTGGCCCTGGCCCTCGAGGGCGGAAGC  
 GGGGTCCGGATGGCCCTCCGCCTCATGGAAGGGGAGTTCC 760  
 CCGACTACCAGAGGGTCATCCCCCAGGAGTTCGCCCTCAA  
 GGTCCAGGTGGAGGGGGAGGCCCTCAGGGAGGCGGTGCGC 840  
 CGGGTGAGCGTCTCTCCGACCGGCAGAACCACCGGGTGG  
 ACCTCCTTTTGGAGGAAGGCCGGATCCTCCTCTCCGCCGA 920  
 GGGGGACTACGGCAAGGGGCAGGAGGAGGTGCCCCGCCAG  
 GTGGAGGGGCGGACATGGCCGTGGCCTACAACGCCCGCT 1000  
 ACCTCCTCGAGGCCCTCGCCCCCGTGGGGGACCGGGCCCA  
 CCTGGGCATCTCCGGGCCCCACGAGCCCGAGCCTCATCTGG 1080  
 GGGGACGGGGAGGGGTACCGGGCGGTGGTGGTGCCCCCTCA  
 GGGTCTAG 1128

FIG.21A

MNITVPKKLLSDQLSLLERIVPSRSANPLYTYLGLYAEEG 40  
ALILFGTNGEVDLEVRLPAEAQSLPRVLVPAQPFFQLVRS  
LPGDLVALGLASEPGQGGQLELSSGRFRTRLAPAEGLYP 120  
- ELLVPEGEDKGAFPLRTRMPSELVKALTHVRYAASNEEY  
RAIFRGVQLEFSPQGFRAVASDGYRLALYDLPLPQGFQAK 200  
AVVPARSVDEMVRVLKGADGAEAVLALGEGVLALALEGGS  
GVRMALRLMEGEFPDYQRVIPQEFALKVQVEGEALREAVR 280  
RVSVLSDRQNHVRVDLLLEEGRILLSAEGDYGKGQEEVPAQ  
VEGPDMAVAYNARYLLEALAPVGDRHLGISGPTSPSLIW 360  
GDGEGYRAVVVPLRVZ

FIG.21B

T.th.beta  
E.coli.bet  
P.mirab.be  
H.infl.bet  
P.put.beta  
B.cap.beta

MNITVPKLLSDQJSLLEIVPSRSANPLYTYLGLYAEEGALLFGTNGEVDLEVRLP  
AE MKFTVEREHLKPLQOVSGPLGGRPTLPILGNLLQVADGTLSLGTDL  
EMEMVARVALV MKFIIEREQLLKPLQOVSGPLGGRPTLPILGNLLKVTENTLSLGTDL  
EMEMMARVSL MQFSISRENLLKPLQOVSGVLSNRPNIPVNNVLLQIEDYRLTITGTDL  
EVELSSQIQLS MHFTIQREALKPLQVAGVVERQTLPVLSNVLVQGGQLSLGTDL  
EVELVGRVQLE MKFTIQNDILTKMLKKITRVLVKNISFPIL  
ENILIQVEDGTLSLTTNLEIELISKIEII

T.th.beta  
E.coli.bet  
P.mirab.be  
H.infl.bet  
P.put.beta  
B.cap.beta

AQSLP-RVLVPAQFFQLVRSIPGLVALGLASEPGGGQLELSSGRFTRLSLAP  
AEGY QPHEPGATTVPARKFFDICRGLP-EGAEIAVQLE---GERMLVRSGRS  
RFSLSLTPAADF QSHEIGATTVPARKFFDIWRGLP-EGAEISVELD---GDRLLVRSGRS  
RFSLSLTPASDF SSSENGTFTIPAKKFIDICRTLS-DDSEITVTFE---QDRALVQSGRS  
RFTLATQPAEEY EPAEPGEITVPARKIMDICKSLP-NDALIDIKVD---EQKLLVKA  
GRSRTLSLTPANDF TKYIPGKTTISGRKIINICRTLS-EKSKIKMQLK---NKKMYISSEN  
SNYILSTLSADTF

T.th.beta  
E.coli.bet  
P.mirab.be  
H.infl.bet  
P.put.beta  
B.cap.beta

PELLVPEGEDKGAPFLTRMPSGELVKALTHVRYAASNEEYRAIFRGVQLEFS  
PQGFRVAV PNLDD---WQSEVEFTLPQAT----MKRLIEATQFSMAHQDVRY  
YLANGMLFETEGEELRTV PNLDD---WQSEVEFTLPQAT----LKRLIEATQFSMAHQDVRY  
YLANGMLFETENTELRTV PNLTD---WQSEVDFELPONT----LRRLIEATQFSMA  
NQDARYFLNGMKFETEGNLLRTV PTVEE---GPGSLTCNLEQSK----LRRLIERTS  
FAMAQQDVRYYLANGMLLEVSRNTLRVAV PNHQN---FDYISKFDISSNI----LKEMIEKTEFS  
MGKQDVRYYLANGMLLEKKDKFLRSV

T.th.beta  
E.coli.bet  
P.mirab.be  
H.infl.bet  
P.put.beta  
B.cap.beta

ASDGYRLALYDLPLPQGFQA--KAVVPARSVDENVRLKGADGAENVLALGEGV  
LALALE ATDGHRLAVCSMPIGQSLPS-HSVIVPRKGVIELMRMLDG-GDNPLRVQIGS  
NNIRAHVG ATDGHRLAVCAMDIGQSLPG-HSVIVPRKGVIELMRLLDGS  
GESLLQLQIGSNNLRAHVG ATDGHRLAVCTISLEQELQN-HSVILPRKGVLELVR  
LLET-NDEPARLQIGTNNLVRVHIK STDGHRLLALCSMSAPIEQEDRHQVIVPRKGI  
LELARLLTD-PEGMVSIVLGQHHIRATTG ATDGYRLAISYTLKKDINF-FSIIIPNKAV  
MELLKLANT-QPQLNILIGSNSIRIYTK

FIG.22A



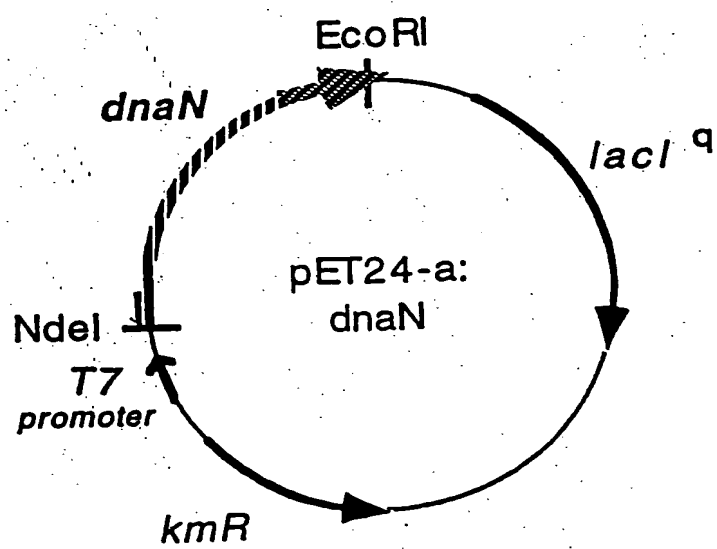
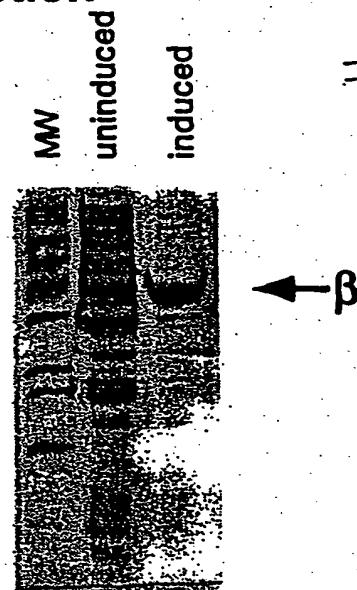


FIG.23

**FIG.24A** Induction



↓  
**Lysis**  
↓  
**Heat Step**  
↓

**FIG.24B** MonoQ Column

Fraction: 5 7 9 11 13 15 17 19 21 23 25

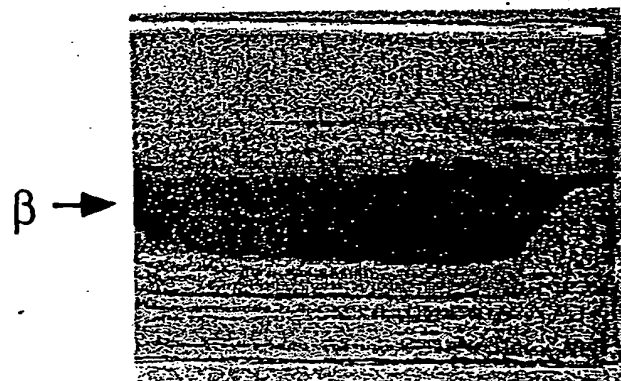


FIG.25A

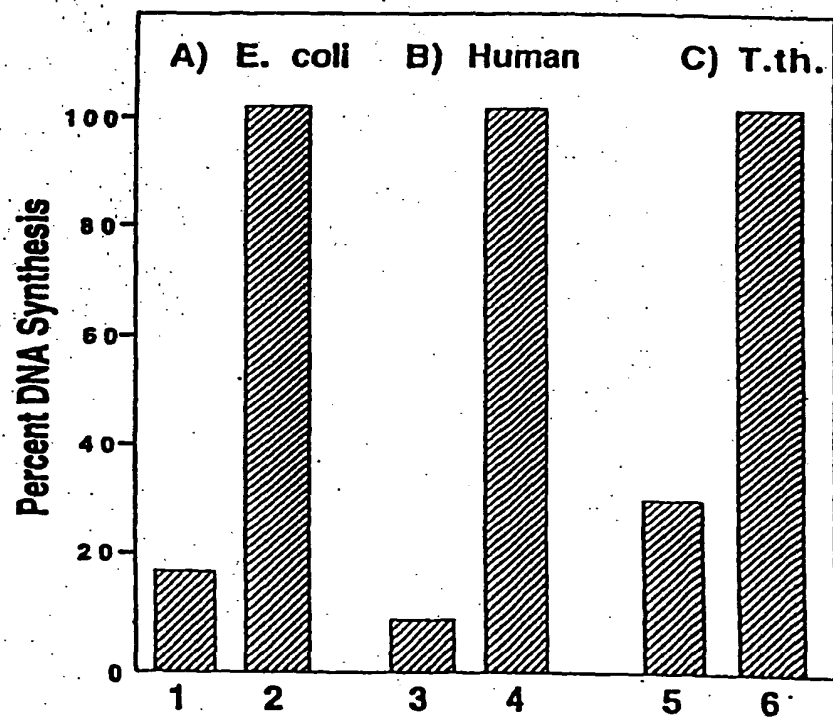
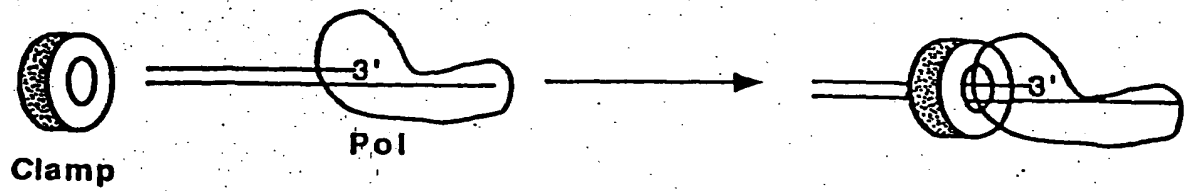


FIG.25B



FIG.26A

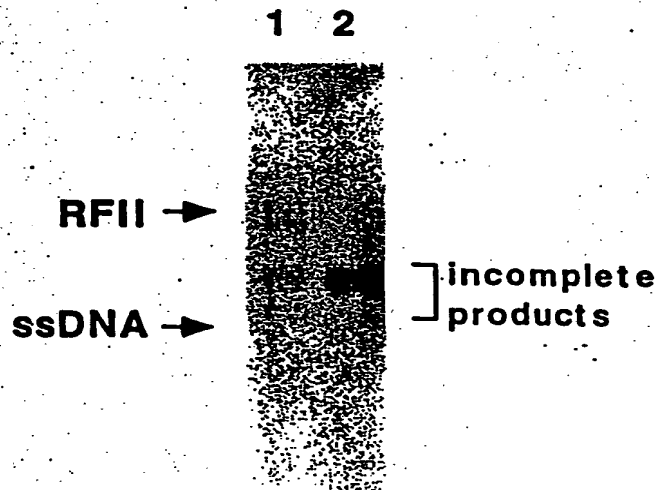
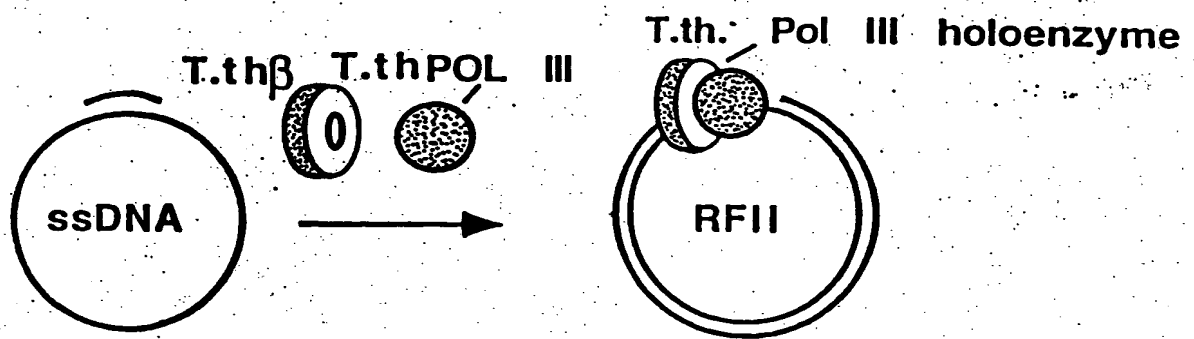


FIG.26B

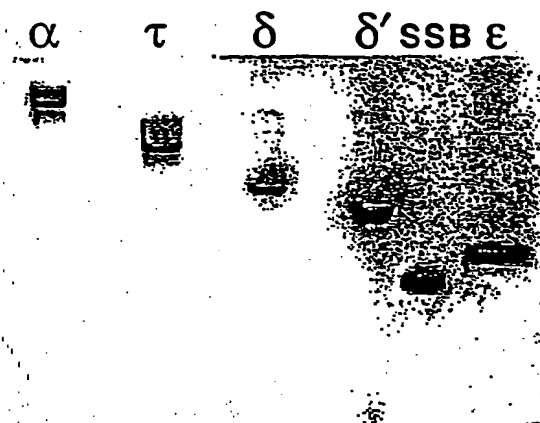


FIG. 27

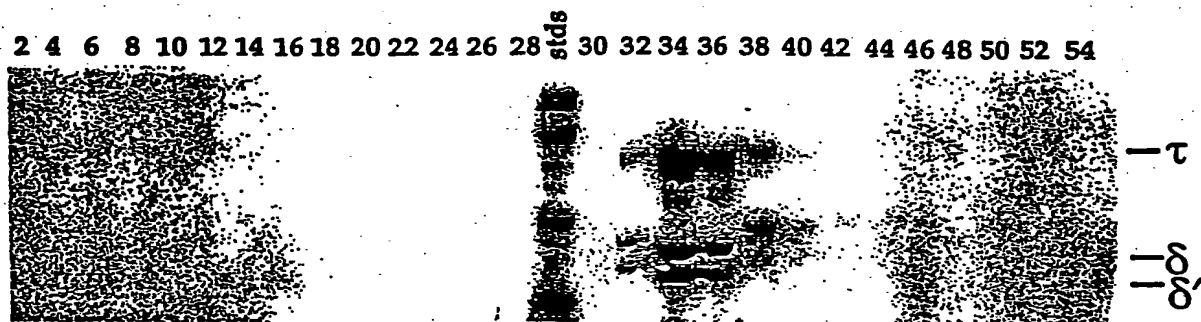


FIG. 28

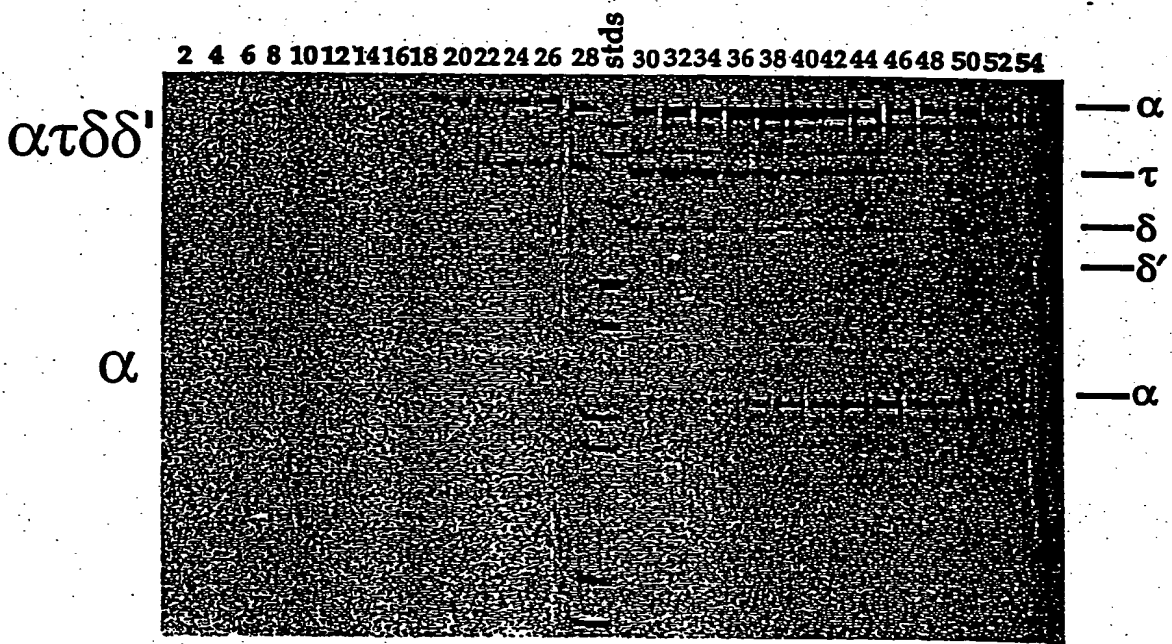


FIG. 29

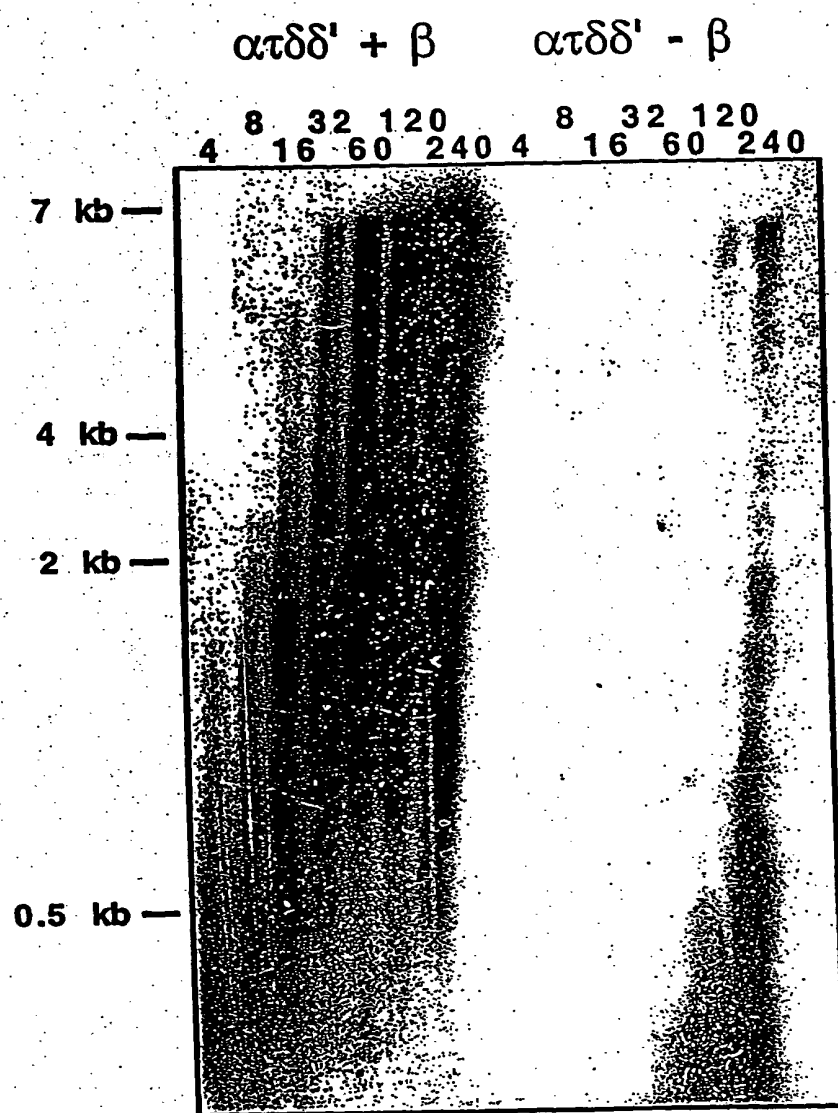
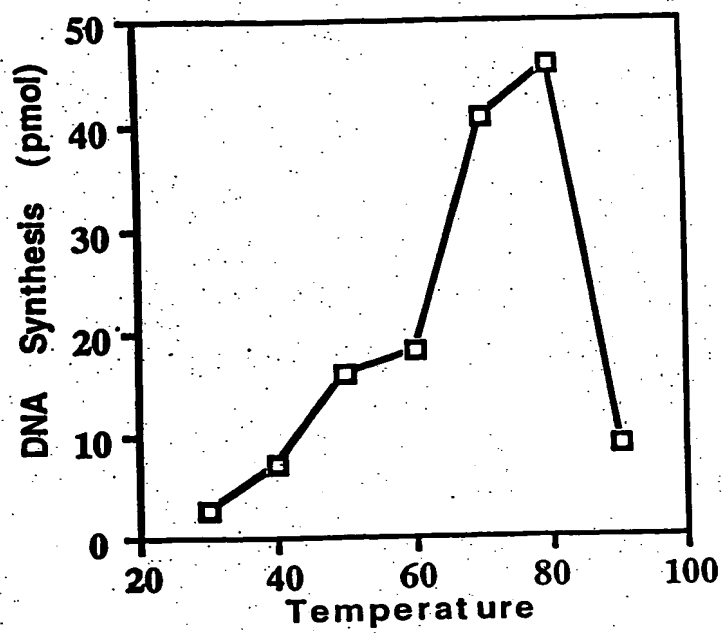
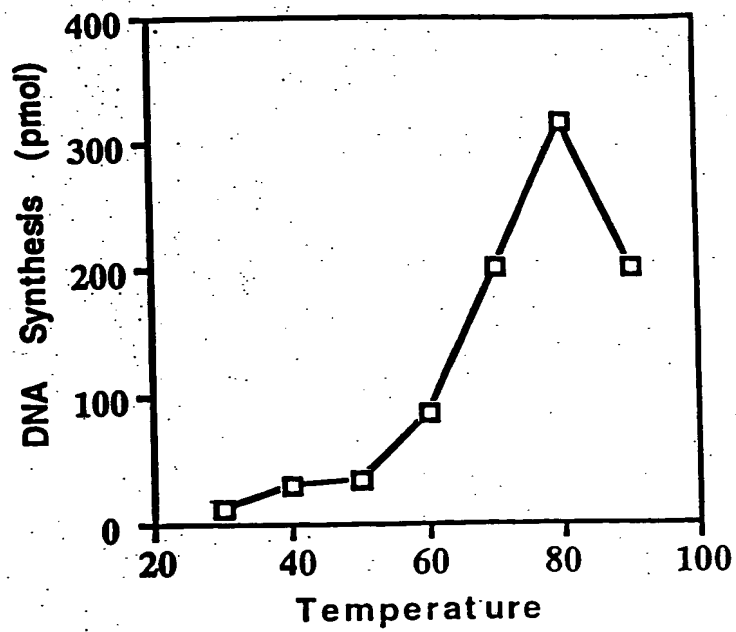


FIG. 30



**FIG. 31**



**FIG. 32**

$\alpha$

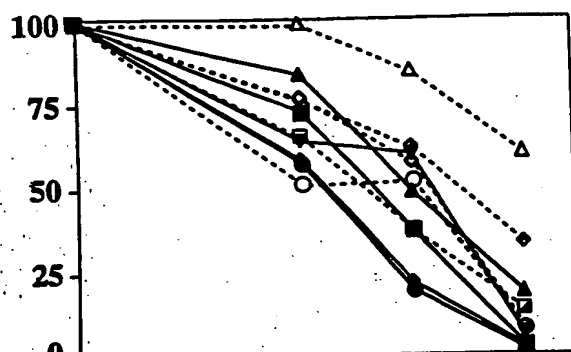


FIG. 33A

$\beta$

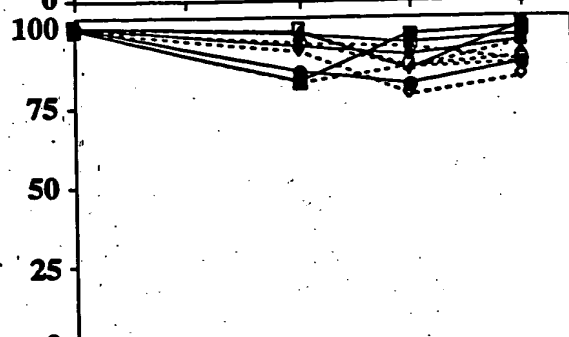


FIG. 33B

$\tau\delta\delta'$

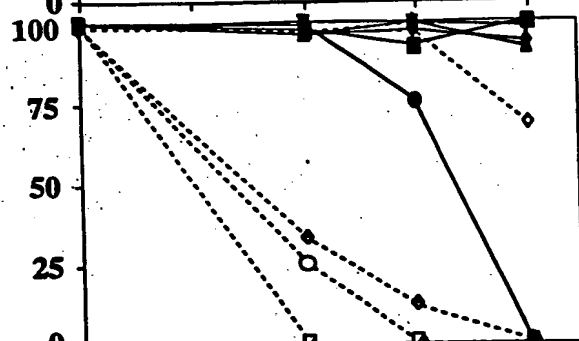


FIG. 33C

SSB

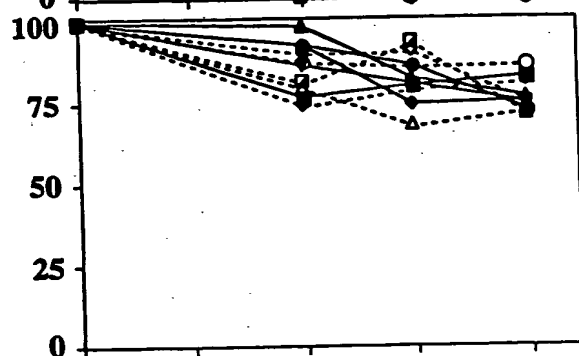


FIG. 33D

Pol III\*

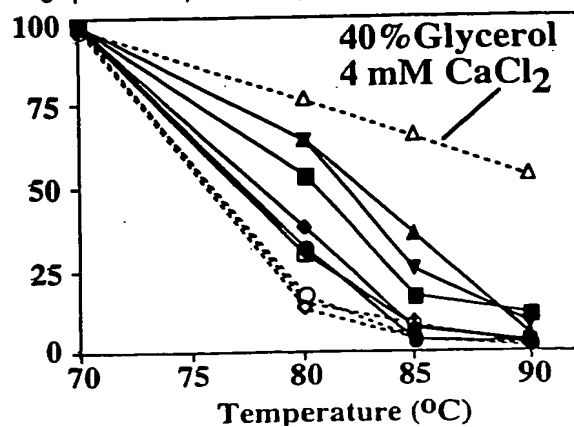


FIG. 33E

|  |      |
|--|------|
| ATGAGTAAGGATTTTCGTCCACCTTCACCTGCACACCCAGTTCTCACTCCT  |      |
| GGACGGGGCTATAAAGATAGACGAGCTCGTGAAAAAGGCAAAGGAGTATG   | 100  |
| GATACAAAGCTGTCTGGAATGTCAGACCACGGAAACCTCTTCGGTTTCGTAT |      |
| AAATTCTACAAAGCCCTGAAGGCGGAAGGAATTAAGCCCATAATCGGCAT   | 200  |
| GGAAGCCTACTTTACCACGGGTTTCGAGGTTTGACAGAAAGACTAAAACGA  |      |
| GCGAGGACAACATAACCGACAAGTACAACCACCACCTCATACTTATAGCA   | 300  |
| AAGGACGAAAAGGTCTAAAGAACTTAATGAAGCTCTCAACCCTCGCCTAC   |      |
| AAAGAAGGTTTTTACTACAAACCCAGAATTGATTACGAACTCCTTGAAAA   | 400  |
| GTACGGGGAGGGCCTAATAGCCCTTACCGCATGCCTGAAAGGTGTTCCCA   |      |
| CCTACTACGCTTCTATAAACGAAGTGAAAAGGCGGAGGAATGGGTAAAG    | 500  |
| AAGTTCAAGGATATATTTCGGAGATGACCTTTATTTAGAACTTCAAGCGAA  |      |
| CAACATTCCAGAACAGGAAGTGCGAAACAGGAACTTAATAGAGATAGCCA   | 600  |
| AAAAGTACGATGTGAAACTCATAGCGACGCAGGACGCCCACTACCTCAAT   |      |
| CCCGAAGACAGGTACGCCACACGGTTCTTATGGCACTTCAAATGAAAAA    | 700  |
| GACCATTACGAACTGAGTTTCGGGAACTTCAAGTGTTCAAACGAAGACC    |      |
| TTCACTTTGCTCCACCCGAGTACATGTGGAAAAAGTTTGAAGGTAAGTTC   | 800  |
| GAAGGCTGGGAAAAGGCACTCCTGAACACTCTCGAGGTAATGGAAAAGAC   |      |
| AGCGGACAGCTTTGAGATATTTGAAAACCTCCACCTACCTCCTTCCCAAGT  | 900  |
| ACGACGTTCCGCCCCGACAAAACCCCTTGAGGAATACCTCAGAGAACTCGCG |      |
| TACAAAGGTTTAAGACAGAGGATAGAAAGGGGACAAGCTAAGGATACTAA   | 1000 |
| AGAGTACTGGGAGAGGCTCGAGTACGAACTGGAAGTTATAAACAAAATGG   |      |
| GCTTTGCGGGATACTTCTTGATAGTTTCAGGACTTCATAAACTGGGCTAAG  | 1100 |
| AAAAACGACATACCTGTTGGACCCGGAAGGGGAAGTGCTGGAGGTTCCCT   |      |
| CGTCGCATACGCCATCGGAATAACGGACGTTGACCCTATAAAGCACGGAT   | 1200 |
| TCCTTTTTTGAGAGGTTCTTAAACCCCGAAAGGGTTTCCATGCCGGATATA  |      |
| GACGTGGATTTCTGTCTCAGGACAACAGGGAAAAGGTCATAGAGTACGTAAG | 1300 |
| GAACAAGTACGGACACGACAACGTAGCTCAGATAATCACCTACAACGTAA   |      |
| TGAAGGCGAAGCAAACACTGAGAGACGTCGCAAGGGCCATGGGACTCCCC   | 1400 |
| TACTCCACCGCGGACAAACTCGCAAAACTCATTCCCTCAGGGGGACGTTCA  |      |
| GGGAACGTGGCTCAGTCTGGAAGAGATGTACAAAACGCCTGTGGAGGAAC   | 1500 |
| TCCTTCAGAAGTACGGAGAACACAGAACGGACATAGAGGACAACGTAAAG   |      |
| AAGTTCAGACAGATATGCGAAGAAAGTCCGGAGATAAAACAGCTCGTTGA   | 1600 |
| GACGGCCCTGAAGCTTGAAGGTCTCACGAGACACACCTCCCTCCACGCCG   |      |
| CGGGAGTGGTTATAGCACCAAAGCCCTTGAGCGAGCTCGTTCCCCTCTAC   | 1700 |
| TACGATAAAGAGGGCGAAGTCGCAACCCAGTACGACATGGTTCAGCTCGA   |      |
| AGAACTCGGTCTCCTGAAGATGGACTTCTCGGACTCAAACCCCTCACAG    | 1800 |
| AACTGAAACTCATGAAAGAACTCATAAAGGAAAGACACGGAGTGATATA    |      |
| AACTTCCTTGAACCTTCCCCCTTGACGACCCGAAAGTTTACAAACTCCTTCA | 1900 |
| GGAAGGAAAAACCACGGGAGTGTTCCAGCTCGAAAGCAGGGGAATGAAAG   |      |
| AACTCCTGAAGAACTAAAGCCCGACAGCTTTGACGACATCGTTGCGGTC    | 2000 |
| CTCGCACTCTACAGACCCGGACCTCTAAAGAGCGGACTCGTTGACACATA   |      |
| CATTAAGAGAAAGCACGGAAAAGAACCCGTTGAGTACCCCTTCCCGGAGC   | 2100 |
| TTGAACCCGTCCTTAAGGAAACCTACGGAGTAATCGTTTATCAGGAACAG   |      |
| GTGATGAAGATGTCTCAGATACTTTCCGGCTTACTCCCGGAGAGGCGGA    | 2200 |
| TACCCTCAGAAAGGCGATAGGTAAGAAGAAAGCGGATTTAATGGCTCAGA   |      |
| TGAAAGACAAGTTCATACAGGGAGCGGTGGAAAGGGGATACCCTGAAGAA   | 2300 |
| AAGATAAGGAAGCTCTGGGAAGACATAGAGAAGTTTCGCTTCCCTACTCCTT |      |
| CAACAAGTCTCACTCGGTAGCTTACGGGTACATCTCCTACTGGACCGCCT   | 2400 |

FIG. 34A

|   |      |
|---|------|
| ACGTTAAAGCCCACTATCCCGCGGAGTTCTTCGCGGTAAAACTCACAAC   |      |
| GAAAAGAACGACAACAAGTTCCTCAACCTCATAAAAGACGCTAAACTCTT  | 2500 |
| CGGATTTGAGATACTTCCCCCGACATAAAACAGAGTGATGTAGGATTTA   |      |
| CGATAGAAGGTGAAAACAGGATAAGGTTCCGGGCTTGCGAGGATAAAGGGA | 2600 |
| GTGGGAGAGGAAACTGCTAAGATAATCGTTGAAGCTAGAAAGAAGTATAA  |      |
| GCAGTTCAAAGGGCTTGCGGACTTCATAAACAAAACCAAGAACAGGAAGA  | 2700 |
| TAAACAAGAAAGTCGTGGAAGCACTCGTAAAGGCAGGGGCTTTTGACTTT  |      |
| ACTAAGAAAAAGAGGAAAGAACTACTCGCTAAAGTGGCAAACTCTGAAAA  | 2800 |
| AGCATTAATGGCTACACAAAACCTCCCTTTTCGGTGCACCGAAAGAAGAAG |      |
| TGGAAGAACTCGACCCCTTAAAGCTTGAAAAGGAAGTTCTCGGTTTTTAC  | 2900 |
| ATTTCAGGGCACCCCTTGACAACCTACGAAAAGCTCCTCAAGAACCGCTA  |      |
| CACACCCATTGAAGATTTAGAAGAGTGGGACAAGGAAAGCGAAGCGGTGC  | 3000 |
| TTACAGGAGTTATCACGGAACCTCAAAGTAAAAAAGACGAAAAACGGAGAT |      |
| TACATGGCGGTCTTCAACCTCGTTGACAAGACGGGACTAATAGAGTGTGT  | 3100 |
| CGTCTTCCCGGGAGTTTACGAAGAGGCAAAGGAACTGATAGAAGAGGACA  |      |
| GAGTAGTGGTAGTCAAAGGTTTTCTGGACGAGGACCTTGAAACGGAAAAT  | 3200 |
| GTCAAGTTCGTGGTGAAAGAGGTTTTCTCCCCTGAGGAGTTCGCAAAGGA  |      |
| GATGAGGAATACCCTTTATATATTCTTAAAAAGAGAGCAAGCCCTAAACG  | 3300 |
| GCGTTGCCGAAAACTAAAGGGAATTATTGAAAACAACAGGACGGAGGAC   |      |
| GGATACAACCTTGGTTCTCACGGTTGATCTGGGAGACTACTTCGTTGATTT | 3400 |
| AGCACTCCCAAGATATGAACTAAAGGCTGACAGAAAGGTTGTAGAGG     |      |
| AGATAGAAAACTGGGAGTGAAGGTCATAATTTAGTAAATAACCCTTACT   | 3500 |
| TCCGAGTAGTCCCC                                      |      |

**FIG. 34B**



|   |      |
|---|------|
| MSKDFVHLHLHTQFSLLDGAIKIDELVKKAKEYGYKAVGMSDHGNLFGSY  |      |
| KFYKALKAEGIKPIIGMEAYFTTGSRFDRKTKTSEDNITDKYNHHLILIA  | 100  |
| KDDKGLKNLMKLSTLAYKEGFYKPRIDYELLEKYGEGLIALTACLKGVF   |      |
| TYYASINEVKKAEWVKFKDIFGDDLYLELQANNIPEQEVANRNLIIEIA   | 200  |
| KKYDVKLIATQDAHYLNPEDRYAHTVLMALQMKKTIHELSSGNFKCSNED  |      |
| LHFAPPEYMWKKFEGKFEGWEKALLNTLEVMEKTADSFEIFENSTYLLPK  | 300  |
| YDVPPDKTLEEYLRELAYKGLRQRIERGQAKDTKEYWERLEYELEVINKM  |      |
| GFAGYFLIVQDFINWAKNDIPVGPGRGSAGGSLVAYAIGITDVPDIKHG   | 400  |
| FLFERFLNPERVSMPTDIDVDFCQDNREKVIIEYVRNKGHDNVAQIITYNV |      |
| MKAKQTLRDVARAMGLPYSTADKLAKLIPQGDVQGTWLSLEEMYKTPVEE  | 500  |
| LLQKYGEHRDIEDNVKKFRQICEESPEIKQLVETALKLEGLTRHTSLHA   |      |
| AGVVIAPKPLSELVPLYDYDKEGEVATQYDMVQLEELGLLKMDFLGLKTLT | 600  |
| ELKLMKELIKERHGVDFINLELPLDDPKVYKLLQEGKTTGVFQLESRGMK  |      |
| ELLKKLKPDSFDDIVAVLALYRPGPLKSGLVDTYIKRKHGKEPVEYFPPE  | 700  |
| LEPVLKETYGVIYVQEQVMKMSQILSGFTPGEADTLRKAIGKKKADLMAQ  |      |
| MKDKFIQGAVERGYPEEKIRKLWEDIEKFASYSFNKSHSVAYGYISYWTA  | 800  |
| YVKAHYPAEFFAVKLTTTEKNDNKFLNLIKDAKLFGFEILPPDINKSDVGF |      |
| TIEGENRIRFGLARIKGVGEETAKIIVEARKKYKQFKGLADFINKTKNRK  | 900  |
| INKKVVEALVKAGAFDFTKKKRKELLAKVANSEKALMATQNSLFGAPKEE  |      |
| VEELDPLKLEKEVLGFYISGHPLDNYEKLKKNRYTPIEDLEEWDKESEAV  | 1000 |
| LTGVITELKVKKTKNGDYMAVFNLVDKTGLIECVVFPGVYEEAKELIEED  |      |
| RVVVVKGFLEDEDLETENVKFVVEKVFSPPEFAKEMRNTLYIFLKREQALN | 1100 |
| GVAEKLKGIENNRTEDGYNLVLTVDLGDYFVDLALPQDMKLKADRKVVE   |      |
| EIEKLGVKVII   | 1161 |

FIG. 35

|  |      |
|--|------|
| ATGAACTACGTTCCCTTCGCGAGAAAGTACAGACCGAAATTCTTCAGGGA   |      |
| AGTAATAGGACAGGAAGCTCCCGTAAGGATACTCAAAAACGCTATAAAAA   | 100  |
| ACGACAGAGTGGCTCACGCCTACCTCTTTGCCGGACCGAGGGGGGTTGGG   |      |
| AAGACGACTATTGCAAGAATTCTCGCAAAGCCTTGAAGTGTA AAAATCC   | 200  |
| CTCAAAGGTGAGCCCTGCGGTGAGTGCGAAAACCTGCAGGGAGATAGACA   |      |
| GGGGTGTGTTCCCTGACTTAATTGAAATGGATGCCGCCTCAAACAGGGGT   | 300  |
| ATAGACGACGTAAGGGCATTAAAAGAAGCGGTCAATTACAAACCTATAAA   |      |
| AGGAAAGTACAAGGTTTACATAATAGACGAAGCTCACATGCTCACGAAAG   | 400  |
| AAGCTTTCAACGCTCTCTTAAAAACCCTCGAAGAGCCCCCTCCAGAACT    |      |
| GTTTTCGTCTTTTGTACCACGGAGTACGACAAAATTCTTCCACGATACT    | 500  |
| CTCAAGGTGTCAGAGGATAATCTTCTCAAAGGTAAGAAAGGAAAAAGTAA   |      |
| TAGAGTATCTAAAAAGATATGTGAAAAGGAAGGGATTGAGTGCGAAGAG    | 600  |
| GGAGCCCTTGAGGTTCTGGCTCATGCCTCTGAAGGGTGCATGAGGGATGC   |      |
| AGCCTCTCTCCTGGACCAGGCGAGCGTTTACGGGGAAGGCAGGGTAACAA   | 700  |
| AAGAAGTAGTGGAGAACTTCTCCTCGGAATTCTCAGTCAGGAAAGCGTTAGG |      |
| AGTTTTCTGAAATTGCTTCTGAACTCAGAAAGTGGACGAAGCTATAAAGTT  | 800  |
| CCTCAGAGAACTCTCAGAAAAGGGCTACAACCTGACCAAGTTTTTGGGAGA  |      |
| TGTTAGAAGAGGAAGTGAGAAACGCAATTTTAGTAAAGAGCCTGAAAAAT   | 900  |
| CCCGAAAGCGTGGTTCAGAACTGGCAGGATTACGAAGACTTCAAAGACTA   |      |
| CCCTCTGGAAGCCCTCCTCTACGTTGAGAACCTGATAAACAGGGGTAAAG   | 1000 |
| TTGAAGCGAGAACGAGAGAACCCTTAAGAGCCTTTGAACTCGCGGTAATA   |      |
| AAGAGCCTTATAGTCAAAGACATAATTCCCGTATCCCAGCTCGGAAGTGT   | 1100 |
| GGTAAAGGAAACCAAAAAGGAAGAAAAGAAAGTTGAAGTAAAAGAAGAGC   |      |
| CAAAAGTAAAAGAAGAAAAACCAAAGGAGCAGGAAGAGGACAGGTTCCAG   | 1200 |
| AAAGTTTTAAACGCTGTGGACGGCAAAATCCTTAAAAGAATACTTGAAGG   |      |
| GGCAAAAAGGGGAAGAAAGAGACGGAAAAATCGTCTTAAAGATAGAAGCCT  | 1300 |
| CTTATCTGAGAACCATGAAAAAGGAATTTGACTCACTAAAGGAGACTTTT   |      |
| CCTTTTTTAGAGTTTGAACCCGTGGAGGATAAAAAAAACCTCAGAAGTC    | 1400 |
| CAGCGGGACGAGGCTGTTTTAAAGGTAAAGGAGCTCTTCAATGCAAAAAT   |      |
| ACTCAAAGTACGAAGTAAAAGCTAAGGTCATAAAGGTGAGAATGCCCGTG   | 1500 |
| GAAGAGATAGGGCTGTTTAAACGCACTAATAGACGGCTTGCCAGGTACGC   |      |
| ACTCACGAGGACGAAGGAAAAGGGAAAGGGAGAAGTTTTCGTTTTAGCGA   | 1600 |
| CTCCTTATAAAGTCAAGGAATTGATGGAAGCTATGGAGGGTATGAAAAAA   |      |
| CACATAAAGGATTTAGAAATCCTCGGAGAGACGGATGAGGATTTAACTTT   | 1700 |
| TTAAAGTATGGGTGTATCTGAGCAAAGGTTTAAAGCTAAAAACAAACCTGA  |      |
| AAACCCGAGGGGACCAGCCGAAAGCCATAAAAAAACTCCTTGAAAACCTA   | 1800 |
| AGGAAAGGCGTAAAAGAACAACACTTCTCGGAGTCACGGGAAGCGGAAA    |      |
| GACTTTTACTCTAGCAAACGTAATAGCGAAGTACAACAAACCAACTCTTG   | 1900 |
| TGGTAGTTCACAACAAAATTCTCGCGGCACAGCTATACAGGGAGTTTAA    |      |
| GAACTATTCCCTGAAAACGCTGTAGAGTACTTTGTCTCTTACTACGACTA   | 2000 |
| TTACCAACCTGAAGCCTACATTCCCGAAAAAGATTTATACATAGAAAAGG   |      |
| ACGCGAGTATAAACGAAAGCTGGAACGTTTCAGACACTCCGCCACGATAT   | 2100 |
| CCGTTCTAGAAAGGAGGGACGTTATAGTAGTTGCTTCAGTTTCTTGATA    |      |
| TACGGACTCGGGAAACCTGAGCACTACGAAAACCTGAGGATAAACTCCA    | 2200 |
| AAGGGGAATAAGACTGAACTTGAGTAAGCTCCTGAGGAAACTCGTTGAGC   |      |
| TAGGATATCAGAGAAATGACTTTGCCATAAAGAGGGCTACCTTCTCGGTT   | 2300 |
| AGGGGAGACGTGGTTGAGATAGTCCCTTCTCACACGGAAGATTACCTCGT   |      |
| GAGGGTAGAGTTCTGGGACGACGAAGTTGAAAGAATACTCCTCATGGACG   | 2400 |
| CTCTGAAC   |      |

FIG. 36

|   |     |
|---|-----|
| MNYVPFARKYRPKFFREVIGQEAPVRILKNAIKNDRVAHAYLFAGPRGVG  |     |
| KTTIARILAKALNCKNPSKGEPGECENCREIDRGVFPDLIEMDAASNRG   | 100 |
| IDDVRLKEAVNYKPIKGKYKVYIIDEAHMLTKEAFNALLKTLEPPPPRT   |     |
| VFVLCTTEYDKILPTILSRCQRIIFSQRKEKVEIYLKKICEKEGIECEE   | 200 |
| GALEVLAHASEGCMRDAASLLDQASVYGEGRVTKEVVENFLGILSQESVR  |     |
| SFLKLLLNSEVDEAIKFLRELSEKGYNLTKFWEMLEEEVRNAILVKSLKN  | 300 |
| PESVVQNWQDYEDFKDYPLEALLYVENLINRGKVEARTREPLRAFELAVI  |     |
| KSLIVKDIIPVSQLGSVVKETKKEKKVEVKEEPKVKEEKPKEQEEDRFQ   | 400 |
| KVLNAVVGKILKRILEGAKREERDGVKIVLKIEASYLRTMKKEFDSLKETF |     |
| PFLEFEPVEDKKKPQKSSGTRLF                             | 473 |

**FIG. 37**

|   |      |
|---|------|
| ATGCGCGTTAAGGTGGACAGGGAGGAGCTTGAAGAGGTTCTTAAAAAAGC  |      |
| AAGAGAAAGCACGGAAAAAAGCCGCACTCCCGATACTCGCGAACTTCT    | 100  |
| TACTCTCCGCAAAAGAGGAAAACTTAATCGTAAGGGCAACGGACTTGGA   |      |
| AACTACCTTGTAGTCTCCGTAAAGGGGGAGGTTGAAGAGGAAGGAGAGGT  | 200  |
| TTGCGTCCACTCTCAAAAACTCTACGATATAGTCAAGAACTTAAATTCCG  |      |
| CTTACGTTTACCTTCATACGGAAGGTGAAAACTCGTCATAACGGGAGGA   | 300  |
| AAGAGTACGTACAAACTTCCGACAGCTCCCGCGGAGGACTTTCCCGAATT  |      |
| TCCAGAAATCGTAGAAGGAGGAGAAACACTTTCGGGAAACCTTCTCGTTA  | 400  |
| ACCGAATAGAAAAGGTAGAGTACGCCATAGCGAAGGAAGAAGCGAACATA  |      |
| GCCCTTCAGGGAATGTATCTGAGAGGATACGAGGACAGAATTCACTTTGT  | 500  |
| GTTCCGGACGGTCACAGGCTTGCACTTTATGAACCTCTACGTAAACATTGA |      |
| AAAGAGTGAAGACGAGTCTTTTGCTTACTTCTCCACTCCCGAGTGGAAAC  | 600  |
| TCGCCGTTAGCTCCTGGAAGGAGAATTCCCGGACTACATGAGTGTCTATCC |      |
| CTGAGGAGTTTTCGGCGGAAGTCTTGTTTGAGACAGAGGAAGTCTTAAAG  | 700  |
| GTTTTAAAGAGGTTGAAGGCTTTAAGCGAAGGAAAAGTTTTTCCCGTGAA  |      |
| GATTACCTTAAGCGAAAACCTTGCCATCTTTGAGTTCGCGGATCCGGAGT  | 800  |
| TCGGAGAAGCGAGAGAGGAAATTGAAGTGGAGTACACGGGAGAGCCCTTT  |      |
| GAGATAGGATTCAACGGAAATACCTTATGGAGGCGCTTGACGCCTACGAC  | 900  |
| AGCGAAAGAGTGTGGTTCAAGTTCACAACCCCGACACGGCCACTTTATT   |      |
| GGAGGCTGAAGATTACGAAAAGGAACCTTACAAGTGCATAATAATGCCGA  | 1000 |
| TGAGGGTGTAGCCATGAAAAAGCTTTAATCTTTTTATTGAGCTTGAGCC   |      |
| TTTTAATTCTGCGTTTAGCGAAGCCAAACCCAAGTCTTC             | 1090 |

FIG. 38

|   |     |
|---|-----|
| MRVKVDREELEEV LKKARESTEKKAALPILANFLLSAKEENLIVRATDLE   |     |
| NYLVVSVKGEVEEEGEV CVHSQKLYDIVKNLNSAYVYLHTEGEKLVITGG   | 100 |
| KSTYKLPTAPAEDFPEFPEIVEGGETLSGNLLVNGIEKVEYAIKEEANI     |     |
| ALQGMYL RGYEDRIHFVGS DGHRLALYEPLGEFSKELLI PRKSLKVLKKL | 200 |
| ITGIEDVNIEKSEDES FAYFSTPEWKLAVRLLEGEFPDYMSVIPEEFSAE   |     |
| VLFEETEEVLKVLKRLKALSEGKVFPVKITLSENLAIFEFADPEFGEAREE   | 300 |
| IEVEYTGEFPEIGFNGKYLMEALDAYDSERVWFKFTTPDTATLLEAEDYE    |     |
| KEPYKCIIMPMRV   | 363 |

FIG. 39

|   |      |
|---|------|
| GTGGAAACCACAATATTCCAGTTCCAGAAAACCTTTTTTCACAAAACCTCC |      |
| GAAGGAGAGGGTCTTCGTCTTCATGGAGAAGAGCAGTATCTCATAAGAA   | 100  |
| CCTTTTTGTCTAAGCTGAAGGAAAAGTACGGGGAGAATTACACGGTTCTG  |      |
| TGGGGGGATGAGATAAGCGAGGAGGAATTCTACACTGCCCTTTCCGAGAC  | 200  |
| CAGTATATTTCGGCGGTTCAAAGGAAAAAGCGGTGGTCATTTACAACCTCG |      |
| GGGATTTCTGAAGAAGCTCGGAAGGAAGAAAAAGGAAAAAGAAAGGCTT   | 300  |
| ATAAAAGTCCTCAGAAACGTAAAGAGTAACTACGTATTTATAGTGACGA   |      |
| TGCGAAACTCCAGAAACAGGAACCTTCTTCGGAACCTCTGAAATCCGTAG  | 400  |
| CGTCTTTTCGGCGGTATAGTGGTAGCAAACAGGCTGAGCAAGGAGAGGATA |      |
| AAACAGCTCGTCTTAAGAAGTTCAAAGAAAAAGGGATAAACGTAGAAAA   | 500  |
| CGATGCCCTTGAATACCTTCTCCAGCTCACGGGTTACAACCTTGATGGAGC |      |
| TCAAACCTTGAGGTTGAAAACTGATAGATTACGCAAGTGAAAAGAAAATT  | 600  |
| TTAACACTCGATGAGGTAAAGAGAGTAGCCTTCTCAGTCTCAGAAAACGT  |      |
| AAACGTATTTGAGTTTCGTTGATTTACTCCTCTTAAAAGATTACGAAAAGG | 700  |
| CTCTTAAAGTTTTGGACTCCCTCATTTCTTCGGAATACACCCCCTCCAG   |      |
| ATTATGAAAATCCTGTCCTCCTATGCTCTAAAACCTTACACCCTCAAGAG  | 800  |
| GCTTGAAGAGAAGGGAGAGGACCTGAATAAGGCGATGGAAAGCGTGGA    |      |
| TAAAGAACAACCTTCTCAAGATGAAGTTCAAATCTTACTTAAAGGCAAAC  | 900  |
| TCTAAAGAGGACTTGAAGAACCTAATCCTCTCCCTCCAGAGGATAGACGC  |      |
| TTTTTCTAAACCTTACTTTTCAGGACACAGTGCAGTTGCTGGGGATTCTT  | 1000 |
| GACCTCAAGACTGGAGAGGGAAGTTGTGAAAAATACTTCTCATGGTGGAT  |      |
| AATCTTTTTTATGAAGTTTGCGGTTTGCGTTTTTCCCGGTTCT         | 1093 |

FIG. 40

|   |     |
|---|-----|
| VETTIFQFKTFFTKPPKERVFLHGEEQYLIRTFLSKLKEYGENYTVL     |     |
| WGDEISEEEFYTALSETSI FGGSKEKAVVIYNFGDFLKKLGRKKKEKERL | 100 |
| IKVLRNVKSNYVFIVYDAKLQKQELSSEPLKSVASFGGIVVANRLSKERI  |     |
| KQLVLKKFKEKGINVENDALEYLLQLTGYNLMELKLEVEKLIDYASEKKI  | 200 |
| LTLDEVKRVAFSVSENVNVFEFVDLLLLLDYKALKVLDLSISFGIHPLO   |     |
| IMKILSSYALKLYTLKRLEEKGEDLNKAMESVG IKNFLKMKFKSYLKN   | 300 |
| SKEDLKNLILSLQRIDAFSKLYFQDTVQLLRDFLTSLREVVKNTSHGG    |     |

FIG. 41

|   |      |
|---|------|
| ATGGAAAAAGTTTTTTTGGAAAACTCCAGAAAACCTTGACATACCCGG    |      |
| AGGACTCCTTTTTTACGGCAAAGAAGGAAGCGGAAAGACGAAAACAGCTT  | 100  |
| TTGAATTTGCAAAAGGTATTTTATGTAAGGAAAACGTACCTGGGGATGCG  |      |
| GAAGTTGTCCCTCCTGCAAACACGTAAACGAGCTGGAGGAAGCCTTCTTT  | 200  |
| AAAGGAGAAATAGAAGACTTTAAAGTTTATAAGACAAGGACGGTAAAAAG  |      |
| CAC TTCGTTTACCTTATGGGCGAACATCCCGACTTTGTGGTAATAATCCC | 300  |
| GAGCGGACATTACATAAAGATAGAACAGATAAGGGAAGTTAAGAACTTTG  |      |
| CCTATGTGAAGCCCGCACTAAGCAGGAGAAAAGTAATTATAATAGACGAC  | 400  |
| GCCACGCGATGACCTCTCAGGCGGCAAACGCTCTTTTAAAGGTATTGGA   |      |
| AGAGCCACCTGCGGACACCACCTTTATCTTGACCACGAACAGGCGTTCTG  | 500  |
| CAATCCTGCCGACTATCCTCTCCAGAACTTTTCAAGTGGAGTTCAAGGGC  |      |
| TTTTCAGTAAAGAGGTTATGGAAATAGCGAAAGTAGACGAGGAAATAGC   | 600  |
| GAAACTCTCTGGAGGCAGTCTAAAAAGGGCTATCTTACTAAAGGAAAACA  |      |
| AAGATATCCTAAACAAAGTAAAGGAATTCTTGAAAACGAGCCGTTAAAA   | 700  |
| GTTTACAAGCTTGCAAGTGAATTCGAAAAGTGGGAACCTGAAAAGCAAAA  |      |
| ACTCTTCCTTGAAATTATGGAAGAATTGGTATCTCAAAAATTGACCGAAG  | 800  |
| AGAAAAAAGACAATTACACCTACCTTCTTGATACGATCAGACTCTTTAAA  |      |
| GACGGACTCGCAAGGGGTGTAAACGAACCTCTGTGGCTGTTTACGTTAGC  | 900  |
| CGTTCAGGCGGATTAATAAACCGTTATTGATTCCGTAACATTTAAACCTT  |      |
| AATCTAAATTATGAGAGCCTTTGAAGGAGGTCTGGTATGGAAAATTTGAA  | 1000 |
| GATTAGATATATAGATACGAGGAAGATAGGAACCGTGAGCGGTGTAAAAG  |      |
| T   | 1051 |

FIG. 42

|  |     |
|--|-----|
| MEKVFLEKLQKTLHIPGGLLFYGKEGSGKTKTAFEFAGKILCKENVPWGC   |     |
| GSCPSCKHVNELEEAFKGEIEDFKVYKDKDGKKHFVYLMGEHPDFVVI     | 100 |
| PSGHYIKIEQIREVKNFAYVKPALSRKVI I IDDAHAMTSQAANALLKVL  |     |
| EEPPADTTFILTTNRRSAILPTILSRTFQVEFKGFSVKEVMEIAKVDEEI   | 200 |
| AKLSGGSCLKRAILLKENKDI LNKVKEFLENEPLKVYKLASEFEKWEPEKQ |     |
| KLFLEIMEELVSQKLTEKKDNYTYLLDTIRLFKDGLARGVNEPLWLFTL    | 300 |
| AVQAD  |     |

FIG. 43

ATGAACTTCCTGAAAAAGTTCCTTTTACTGAGAAAAGCTCAAAAGTCTCC  
 TTA CTTCGAAGAGTTCTACGAAGAAATCGATTTGAACCAGAAGGTGAAAG 100  
 ATGCAAGGTTTGTAGTTTTTGA CTGCGAAGCCACAGAACTCGACGTAAAG  
 AAGGCAAAACTCCTTTCAATAGGTGCGGTTGAGGTTAAAAACCTGGAAAT 200  
 AGACCTCTCTAAATCTTTTTACGAGATACTCAAAAGTGACGAGATAAAGG  
 CGGCGGAGATACATGGAATAACCAGGGAAGACGTTGAAAAGTACGGAAAG 300  
 GAACCAAAGGAAGTAATATACGACTTTCTGAAGTACATAAAGGGAAGCGT  
 TCTCGTTGGCTACTACGTGAAGTTTGACGTCTCACTCGTTGAGAAGTACT 400  
 CCATAAAGTACTTCCAGTATCCAATCATCAACTACAAGTTAGACCTGTTT  
 AGTTTTCGTGAAGAGAGAGTACCAGAGTGGCAGGAGTCTTGACGACCTTAT 500  
 GAAGGAACTCGGTGTAGAAATAAGGGCAAGGCACAACGCCCTTGAAGATG  
 CCTACATAACCGCTCTTCTTTTCCTAAAGTACGTTTACCCGAACAGGGAG 600  
 TACAGACTAAAGGATCTCCCGATTTTCCTT

FIG. 44

MNFLKKFLLLRKAQKSPYFEEFYEEIDLNQKVKDARFVVFDCATELDVK  
 KAKLLSIGAVEVKNLEIDLKSFYEILKSDEIKAAEIHGITREDVEKYGK 100  
 EPKEVIYDFLKYIKGSVLVGYYVKFDVSLVEKYSIKYFQYPIINYLKDLF  
 SFVKREYQSGRSLDDLMKELGVEIRARHNALEDAYITALLFLKYVYPNRE 200  
 YRLKDLPIFL

FIG. 45

|  |     |
|--|-----|
| ATGCTCAATAAGGTTTTTATAATAGGAAGACTTACGGGTGACCCCGTTAT |     |
| AACTTATCTACCGAGCGGAACGCCCCTAGTAGAGTTTACTCTGGCTTACA | 100 |
| ACAGAAGGTATAAAAACAGAACGGTGAATTTTCAGGAGGAAAGTCACTTC |     |
| TTTGACGTAAAGGCGTACGGAAAAATGGCTGAAGACTGGGCTACACGCTT | 200 |
| CTCGAAAGGATACCTCGTACTCGTAGAGGGAAGACTCTCCCAGGAAAAGT |     |
| GGGAGAAAGAAGGAAAGAAGTTCTCAAAGGTCAGGATAATAGCGGAAAAC | 300 |
| GTAAGATTAATAAACAGGCCGAAAGGTGCTGAACTTCAAGCAGAAGAAGA |     |
| GGAGGAAGTTCCTCCATTGAGGAGGAAATTGAAAACTCGGTAAAGAGG   | 400 |
| AAGAGAAGCCTTTTACCGATGAAGAGGACGAAATACCTTTTTTAATTTGA |     |
| GGAGGTTAAAGTATGGTAGTGAGAGCTCCTAAGAAGAAAGTTTGTATGTA | 500 |
| CTGTGAACAAAAGAGAGAGCCAGATT                         |     |

**FIG. 46**

|  |     |
|--|-----|
| MLNKVFIIGRLTGDPVITYLPSGTPVVEFTLAYNRRYKNQNGEFQEESHF |     |
| FDVKAYGKMAEDWATRFSGYLVLEGRLSQEKWEKEGKKFSKVRIIAEN   | 100 |
| VRLINRPKGAEELQAEIEEEVPPIEEEIEKLGKEEEKPFTDEEDEIPF   |     |

**FIG. 47**



|  |      |
|--|------|
| ATGCAATTTGTGGATAAACTTCCCTGTGACGAATCCGCCGAGAGGGCGGT   |      |
| TCTTGGCAGTATGCTTGAAGACCCCGAAAACATACCTCTGGTACTTGAAT   | 100  |
| ACCTTAAAGAAGAAGACTTCTGCATAGACGAGCACAAAGCTACTTTTCAGG  |      |
| GTTCTTACAAACCTCTGGTCCGAGTACGGCAATAAGCTCGATTTTCGTATT  | 200  |
| AATAAAGGATCACCTTGAAAAGAAAACTTACTCCAGAAAATACCTATAG    |      |
| ACTGGCTCGAAGAACTCTACGAGGAGGCGGTATCCCCTGACACGCTTGAG   | 300  |
| GAAGTCTGCAAAATAGTAAAACAACGTTCCGCACAGAGGGCGATAATTCA   |      |
| ACTCGGTATAGAACTCATTACAAAGGAAAGGAAAACAAAGACTTTTACA    | 400  |
| CATTAATCGAGGAAGCCCAGAGCAGGATATTTTCCATAGCGGAAAGTGCT   |      |
| ACATCTACGCAGTTTTTACCATGTGAAAGACGTTGCGGAAGAAGTTATAGA  | 500  |
| ACTCATTTATAAATTCAAAGCTCTGACAGGCTAGTCACGGGACTCCCAA    |      |
| GCGGTTTTCACGGAACCTCGATCTAAAGACGACGGGATTCCACCCTGGAGAC | 600  |
| TTAATAATACTCGCCGCAAGACCCGGTATGGGGAAAACCGCCTTTATGCT   |      |
| CTCCATAATCTACAATCTCGCAAAGACGAGGGAAAACCTCAGCTGTAT     | 700  |
| TTTCCTTGGAATGAGCAAGGAACAGCTCGTTATGAGACTCCTCTCTATG    |      |
| ATGTCGGAGGTCCCACTTTTCAAGATAAGGTCTGGAAGTATATCGAATGA   | 800  |
| AGATTTAAAGAAGCTTGAAGCAAGCGCAATAGAACTCGCAAAGTACGACA   |      |
| TATACCTCGACGACACACCCGCTCTCACTACAACGGATTTAAGGATAAGG   | 900  |
| GCAAGAAAGCTCAGAAAGGAAAAGGAAGTTGAGTTCGTGGCGGTGGACTA   |      |
| CTTGCAACTTCTGAGACCGCCAGTCCGAAAGAGTTCAAGACAGGAGGAAG   | 1000 |
| TGGCAGAGGTTTCAAGAACTTAAAAGCCCTTGCAAAGGAACTTCACATT    |      |
| CCCGTTATGGCACTTGCGCAGCTCTCCCGTGAGGTGGAAAAGAGGAGTGA   | 1100 |
| TAAAAGACCCAGCTTGCGGACCTCAGAGAATCCGGACAGATAGAACAGG    |      |
| ACGCAGACCTAATCCTTTTCTCCACAGACCCGAGTACTACAAGAAAAAG    | 1200 |
| CCAAATCCCGAAGAGCAGGGTATAGCGGAAGTGATAATAGCCAAGCAAAG   |      |
| GCAAGGACCCACGGACATTGTGAAGCTCGCATTATTATAAGGAGTACACTA  | 1300 |
| AGTTTGCAAACCTAGAAGCCCTTCTGAACAACCTCCTGAAGAAGAGGAA    |      |
| CTTTCCGAAATTATTGAAACACAGGAGGATGAAGGATTGAAGATATTGA    | 1400 |
| CTTCTGAAAATTAAGGTTTTATAATTTTATCTTGGCTATCCGGGGTAGCT   |      |
| CAATCGGCAGAGCGGGTGGCTG                               | 1472 |

FIG. 48

|   |     |
|---|-----|
| MQFVDKLPCEDESAERAVLGSMLEDPENIPLVLEYLKEEDFCIDEHKLLFR |     |
| VLTNLWSEYGNKLDFVLIKDHLEKKNLLQKIPIDWLEELYEEAVSPDTLE  | 100 |
| EVCKIVKQRSAQRAIIQLGITSTQFYHVKDVAEEVIELIYKFKSSDRLVT  |     |
| GLPSGFTELDLKTTFHFGDLIILAARPGMGKTAFLMSIIYNLAKDEGKP   | 200 |
| SAVFSLEMSKEQLVMRLLSMMSEVPLFKIRSGSISNEDLKKLEASAIELA  |     |
| KYDIYLLDDTPALTTTDLRIRARKLRKEKEVEFVAVDYLQLLRPPVRKSSR | 300 |
| QEEVAEVSRLKALAKELHIPVMALAQLSREVEKRS DKRPQLADLRESGQ  |     |
| IEQDADLILFLHRPEYKPKPNPEEQGIAEVIIAKQRQGPTDIVKLAFIK   | 400 |
| EYTKFANLEALPEQPPEEEELSEIIETQEDEGFEDIDF              |     |

FIG. 49

|  |      |
|--|------|
| ATGTCCTCGGACATAGACGAACTTAGACGGGAAATAGATATAGTAGACGT   |      |
| CATTTCCGAATACTTAAACTTAGAGAAGGTAGGTTCCAATTACAGAACGA   | 100  |
| ACTGTCCCTTTTACCCTGACGATACACCCTCCTTTTACGTGTCTCCAAGT   |      |
| AAACAAATATTCAAGTGTTCGGTTGCGGGGTAGGGGAGACGCGATAAA     | 200  |
| GTTTCGTTTCCCTTTACGAGGACATCTCCTATTTTGAAGCCGCCCTTGAAC  |      |
| TCGCAAAACGCTACGGAAAGAAATTAGACCTTGAAAAGATATCAAAAGAC   | 300  |
| GAAAAGGTATACGTGGCTCTTGACAGGGTTTGTGATTTCTACAGGGAAAG   |      |
| CCTTCTCAAAAACAGAGAGGCAAGTGAGTACGTAAAGAGTAGGGGAATAG   | 400  |
| ACCCTAAAGTAGCGAGGAAGTTTGATCTTGGGTACGCACCTTCCAGTGAA   |      |
| GCACTCGTAAAAGTCTTAAAAGAGAACGATCTTTTAGAGGCTTACCTTGA   | 500  |
| AACTAAAAACCTCCTTTCTCCTACGAAGGGTGTTCACAGGGATCTCTTTC   |      |
| TTTCGGCGTGTTCGTGATCCCGATAAAGGATCCGAGGGGAAGAGTTATAGGT | 600  |
| TTTCGGTGAAGGAGGATAGTAGAGGACAAATCTCCCAAGTACATAAACTC   |      |
| TCCAGACAGCAGGGTATTTAAAAAGGGGGAGAACTTATTCGGTCTTTACG   | 700  |
| AGGCCAAAGGAGTATATAAAGGAAGAAGGATTTGCGATACTTGTGGAAGGG  |      |
| TACTTTGACCTTTTGAGACTTTTTTCCGAGGGGAATAAGGAACGTTGTTGC  | 800  |
| ACCCCTCGGTACAGCCCTGACCCAAAATCAGGCAAACCTCCTTTCCAAGT   |      |
| TCACAAAAAAGGTCTACATCCTTTACGACGGAGATGATGCGGGAAGAAAG   | 900  |
| GCTATGAAAAGTGCCATTCCCCTACTCCTCAGTGCAGGAGTGGAAGTTTA   |      |
| TCCCGTTTACCTCCCCGAAGGATACGATCCCGACGAGTTTATAAAGGAAT   | 1000 |
| TCGGGAAAGAGGAATTAAGAAGACTGATAAACAGCTCAGGGGAGCTCTTT   |      |
| GAAACGCTCATAAAAACCGCAAGGGAAAACCTTAGAGGAGAAAACGCGTGA  | 1100 |
| GTTTCAGGTATTATCTGGGCTTTATTTCCGATGGAGTAAGGCGCTTTGCTC  |      |
| TGGCTTCGGAGTTTCACACCAAGTACAAAGTTCCTATGGAAATTTTATTA   | 1200 |
| ATGAAAATTGAAAAAAATTCTCAAGAAAAAGAAATTAACTCTCCTTTAA    |      |
| GGAAAAAATCTTCTGAAAGGACTGATAGAATTAAAACCAAAAATAGACC    | 1300 |
| TTGAAGTCCTGAACTTAAGTCCTGAGTTAAAGGAACTCGCAGTTAACGCC   |      |
| TTAAACGGAGAGGAGCATTACTTCCAAAAGAAGTTCTCGAGTACCAGGT    | 1400 |
| GGATAACTTGGAGAAACTTTTTTAACAACATCCTTAGGGATTACAAAAAT   |      |
| CTGGGAAAAAGAGGAAGAAAAGAGGGTTGAAAAATGTAAATACTTAATTA   | 1500 |
| ACTTTAATAAATTTTATAGAGTTAGGA                          |      |

FIG. 50

|   |     |
|---|-----|
| MSSDIDELRREIDIVDVISEYLNLEKVGSNYRTNCPFHPDDTPSFYVSPS  |     |
| KQIFKFCGCGVGDAIKFVSLYEDISYFEAALELAKRYGKKLDLEKISKD   | 100 |
| EKVYVALDRVCDFYRESLLKNREASEYVKSARGIDPKVARKFDLGYAPSSE |     |
| ALVKVLKENDLLEAYLETKNLLSPTKGVYRDLFLRRVVIPIKDPRGRVIG  | 200 |
| FGRRIVEDKSPKYINSPDSRVFKKGENLFLGLYEKEYIKEEGFAILVEG   |     |
| YFDLLRLRFSEGIRNVVAPLGTALTQNQANLLSKFTKKVYILYDGDDAGRK | 300 |
| AMKSAIPLLLSAGVEVYPVYLPEGYDPDEFIKEFGKEELRRLINSSGELF  |     |
| ETLIKTARENLEEKTRFRYYLGFISDGVRRFALASEFHTKYKVPMEILL   | 400 |
| MKIEKNSQEKEIKLSFKEKIFLKGLIELKPKIDLEVLNLSPELKELAVNA  |     |
| LNGEEHLLPKEVLEYQVDNLEKLFNNILRDLQKSGKKRKKRGLKNVNT    | 498 |

FIG. 51

|   |     |
|---|-----|
| ATGCAAGATACCGCTACCTGCAGTATTTGTCAGGGGACGGGATTTCGTAAA |     |
| GACCGAAGACAACAAGGTAAGGCTCTGCGAATGCAGGTTCAAGAAAAGGG  | 100 |
| ATGTAAACAGGGAACTAAACATCCCAAAGAGGTACTGGAACGCCAACTTA  |     |
| GACACTTACCACCCCAAGAACGTATCCAGAACAGGGCACTTTTGACGAT   | 200 |
| AAGGGTCTTCGTCCACAACCTTCAATCCCGAGGAAGGGAAAGGGCTTACCT |     |
| TTGTAGGATCTCCTGGAGTCGGCAAACTCACCTTGCGGTTGCAACATTA   | 300 |
| AAAGCGATTTATGAGAAGAAGGGAATCAGAGGATACTTCTTCGATACGAA  |     |
| GGATCTAATATTCAGGTTAAAACACTTAATGGACGAGGGAAAGGATACAA  | 400 |
| AGTTTTTAAAACTGTCTTAAACTCACCGGTTTTGGTTCTCGACGACCTC   |     |
| GGTTCTGAGAGGCTCAGTGACTGGCAGAGGGAACATCTCTTACATAAT    | 500 |
| CACTTACAGGTATAACAACCTTAAGAGCACGATAATAACCACGAATTACT  |     |
| CACTCCAGAGGGAAGAAGAGAGTAGCGTGAGGATAAGTGCGGATCTTGCA  | 600 |
| AGCAGACTCGGAGAAAACGTAGTTTCAAAAATTTACGAGATGAACGAGTT  |     |
| GCTCGTTATAAAGGGTTCGGACCTCAGGAAGTCTAAAAAGCTATCAACCC  | 700 |
| CATCT   |     |

FIG. 52

|   |     |
|---|-----|
| MQDTATCSICQGTGFVKTEDNKNVRLCECRFKKRDVNRELNIPKRYWNANL |     |
| DTYHPKNVSONRALLTIRVFVHNFNPEEGKGLTFVGSPGVGKTHLAVATL  | 100 |
| KAIYEKKGIRGYFFDTKDLIFRLKHLMDDEGKDTKFLKTVLNSPVLVLDL  |     |
| GSERLSDWQRELISYIITYRYNNLKSTIITNYSLOREEESSVRISADLA   | 200 |
| SRLGENVVSKIYEMNELLVIKGSDLRKS KKLSTPS                |     |

FIG. 53

|  |      |
|--|------|
| ATGAAAAAGATTGAAAATTTGAAGTGGAAAAATGTCTCGTTTAAAAGCCT   |      |
| GGAAATAGATCCCGATGCAGGTGTGGTTCTCGTTTCCGTGGAAAAATTCT   | 100  |
| CCGAAGAGATAGAAGACCTTGTGCGTTTACTGGAGAAGAAGACGCGGTTT   |      |
| CGAGTCATCGTGAACGGTGTTCAAAAAAGTAACGGGGATCTAAGGGGAAA   | 200  |
| GATACTTTCCCTTCTCAACGGTAATGTGCCTTACATAAAAGATGTTGTTT   |      |
| TCGAAGGAAACAGGCTGATTCTGAAAGTGCTTGGAGATTTCGCGCGGGAC   | 300  |
| AGGATCGCCTCCAACTCAGAAGCACGAAAAAACAGCTCGATGAACTGCT    |      |
| GCCTCCCGGAACAGAGATCATGCTGGAGGTGTGGAGCCTCCGGAAGATC    | 400  |
| TTTTGAAAAAGGAAGTACCACAACCAGAAAAGAGAGAAGAACCAAAGGT    |      |
| GAAGAATTGAAGATCGAGGATGAAAACCACATCTTTGGACAGAAACCCAG   | 500  |
| AAAGATCGTCTTCACCCCTCAAAAATCTTTGAGTACAACAAAAAGACAT    |      |
| CGGTGAAGGGCAAGATCTTCAAAATAGAGAAGATCGAGGGGAAAAGAACG   | 600  |
| GTCTTCTGATTTACCTGACAGACGGAGAAGATTCTCTGATCTGCAAAGT    |      |
| CTTCAACGACGTTGAAAAGGTCTGAAGGGGAAAGTATCGGTGGGAGACGTGA | 700  |
| TCGTTGCCACAGGAGACCTCCTTCTCGAAAACGGGGAGCCCACCCTTTAC   |      |
| GTGAAGGGAATCACAAACTTCCCGAAGCGAAAAGGATGGACAAATCTCC    | 800  |
| GGTTAAGAGGGTGGAGCTCCACGCCCATACCAAGTTCAGCGATCAGGACG   |      |
| CAATAACAGATGTGAACGAATATGTGAAACGAGCCAAGGAATGGGGCTTT   | 900  |
| CCCGCGATAGCCCTCACGGATCATGGGAACGTTCAAGCCATACCTTACTT   |      |
| CTACGACGCGGCGAAAGAAGCTGGAATAAAGCCCATTTTCGGTATCGAAG   | 1000 |
| CGTATCTGGTGAAGTACGTGGAGCCCGTCATAAGGAATCTCTCCGACGAT   |      |
| TCGACGTTTGGAGATGCCACGTTCTGTCGTCCTCGACTTCGAGACGACGGG  | 1100 |
| TCTCGACCCGCAGGTGGATGAGATCATCGAGATAGGAGCGGTGAAGATAC   |      |
| AGGGTGGCCAGATAGTGGACGAGTACCACACTCTCATAAAGCCTTCCAGG   | 1200 |
| GAGATCTCAAGAAAAAGTTCGGAGATCACCGGAATCACTCAAGAGATGCT   |      |
| GGAAAACAAGAGAAGCATCGAGGAAGTTCTGCCGGAGTTCTCGGTTTTTC   | 1300 |
| TGGAAGATTCCATCATCGTAGCACACAACGCCAACTTCGACTACAGATTT   |      |
| CTGAGGCTGTGGATCAAAAAAGTGATGGGATTGGACTGGGAAAGACCCTA   | 1400 |
| CATAGATACGCTCGCCCTCGCAAAGTCCCTTCTCAAACCTGAGAAGCTACT  |      |
| CTCTGGATTCCGTTGTGGAAAAGCTCGGATTGGGTCCCTTCCGGCACCAC   | 1500 |
| AGGGCCCTGGATGACGCGAGGGTCAACGCTCAGGTTTTCTCAGGTTCTGT   |      |
| TGAGATGATGAAGAAGATCGGTATCACGAAGCTTTCAGAAATGGAGAAGT   | 1600 |
| TGAAGGATACGATAGACTACACCGCGTTGAAACCTTCCACTGCACGATC    |      |
| CTCGTTCAGAACAAAAAGGGATTGAAAAACCTATACAAACTGGTTTCTGA   | 1700 |
| TTCTATATAAAGTACTTCTACGGTGTTCCGAGGATCCTCAAAAGTGAGC    |      |
| TCATCGAGAACAGAGAAGGACTGCTCGTGGGTAGCGCGTGTATCTCCGGT   | 1800 |
| GAGCTCGGACGTGCCGCCCTCGAAGGAGCGAGTGATTCAGAACTCGAAGA   |      |
| GATCGCGAAGTTCTACGACTACATAGAAGTCATGCCGCTCGACGTTATAG   | 1900 |
| CCGAAGATGAAGAAGACCTAGACAGAGAAAGACTGAAAGAAGTGTACCGA   |      |
| AAACTCTACAGAATAGCGAAAAAATTGAACAAGTTCGTCGTCATGACCGG   | 2000 |
| TGATGTTTCAATTCCTCGATCCCAGAGATGCCAGGGGCAGAGCTGCACTTC  |      |
| TGGCACCTCAGGGAAACAGAACTTCGAGAATCAGCCCGCACTCTACCTC    | 2100 |
| AGAACGACCGAAGAAATGCTCGAGAAGGCGATAGAGATATTCTGAAGATGA  |      |
| AGAGATCGCGAGGGAAGTCGTGATAGAGAATCCCAACAGAATAGCCGATA   | 2200 |
| TGATCGAGGAAGTGCAGCCGCTCGAGAAAAAATTCACCCGCCGATCATA    |      |
| GAGAACGCCGATGAAATAGTGAGAAACCTCACCATGAAGCGGGCGTACGA   | 2300 |
| GATCTACGGTGATCCGCTTCCCGAAATCGTCCAGAAGCGTGTGGAAAAGG   |      |

FIG. 54A

|  |      |
|--|------|
| AACTGAACGCCATCATAAATCATGGATACGCCGTTCTCTATCTCATCGCT   | 2400 |
| CAGGAGCTCGTTTCAGAAATCTATGAGCGATGGTTACGTGGTTGGATCCAG  |      |
| AGGATCCGTCGGGTCTTCACTCGTGGCCAATCTCCTCGGAATAACAGAGG   | 2500 |
| TGAATCCCCTACCACCACATTACAGGTGTCCAGAGTGCAAATACTTTGAA   |      |
| GTTGTCTGAAGACGACAGATACGGAGCGGGTTACGACCTTCCCAACAAGAA  | 2600 |
| CTGTCCAAGATGTGGGGCTCCTCTCAGAAAAGACGGCCACGGCATAACCGT  |      |
| TTGAAACGTTTCATGGGGTTCGAGGGTGACAAGGTCCCCGACATAGATCTC  | 2700 |
| AACTTCTCAGGAGAGTATCAGGAACGTGCTCATCGTTTTGTGGAAGAACT   |      |
| CTTCGGTAAAGACCACGTCCTATAGGGCGGGAACCATAAACACCATCGCGG  | 2800 |
| AAAGAAGTGCGGTGGGTACGTGAGAAGCTACGAAGAGAAAACCGGAAAG    |      |
| AAGCTCAGAAAGGCGGAAATGGAAAGACTCGTTTCCATGATCACGGGAGT   | 2900 |
| GAAGAGAACGACGGGTACGACCCAGGGGGGCTCATGATCATACCGAAAG    |      |
| ACAAAGAAGTCTACGATTTCACTCCCATACAGTATCCAGCCAACGATAGA   | 3000 |
| AACGCAGGTGTGTTCAACACGCACTTCGCATACGAGACGATCCATGATGA   |      |
| CCTGGTGAAGATAGATGCGCTCGGCCACGATGATCCCACTTTCATCAAGA   | 3100 |
| TGCTCAAGGACCTCACCGGAATCGATCCCATGACGATTCCCATGGATGAC   |      |
| CCCGATACGCTCGCCATATTAGTTCTGTGAAGCCTCTTGGTGTGGATCC    | 3200 |
| CGTTGAGCTGGAAAGCGATGTGGGAACGTACGGAATTCCGGAGTTCGGAA   |      |
| CCGAGTTTGTGAGGGGAATGCTCGTTGAAACGAGACCAAAGAGTTTCGCC   | 3300 |
| GAGCTTGTGAGAATCTCAGGACTGTACACGGTACGGACGTCTGGTTGAA    |      |
| CAACGCACGTGATTGGATAAACCTCGGCTACGCCAAGCTCTCCGAGGTTA   | 3400 |
| TCTCGTGTAGGGACGACATCATGAACTTCCTCATACACAAAGGAATGGAA   |      |
| CCGTCACCTTGCCTTCAAGATCATGGAAAACGTACAGGAAGGGAAAGGGTAT | 3500 |
| CACAGAAGAGATGGAGAGCGAGATGAGAAGGCTGAAGGTTCCAGAATGGT   |      |
| TCATCGAATCCTGTAAAAGGATCAAATATCTCTTCCCGAAAGCTCACGCT   | 3600 |
| GTGGCTTACGTGAGTATGGCCTTCAGAATTGCTTACTTCAAGGTTCACTA   |      |
| TCCTCTTCAGTTTTTACGCGGCGTACTTCACGATAAAAGGTGATCAGTTCG  | 3700 |
| ATCCGGTTCTCGTACTCAGGGGAAAAGAAGCCATAAAGAGGCGCTTGAGA   |      |
| GAACTCAAAGCGATGCCTGCCAAAGACGCCCAGAAGAAAAACGAAGTGAG   | 3800 |
| TGTTCTGGAGGTTGCCCTGGAAATGATACTGAGAGGTTTTTCTTCTCTAC   |      |
| CGCCCGACATCTTCAAATCCGACGCGAAGAAATTTCTGATAGAAGGAAAC   | 3900 |
| TCGCTGAGAATTCCGTTCAACAAACTTCCAGGACTGGGTGACAGCGTTGC   |      |
| CGAGTCGATAATCAGAGCCAGGGAAGAAAAGCCGTTCACTTCGGTGGAAG   | 4000 |
| ATCTCATGAAGAGGACCAAGGTCAACAAAAATCACATAGAGCTGATGAAA   |      |
| AGCCTGGGTGTTCTCGGGGACCTTCCAGAGACGGAACAGTTCACGCTTTT   | 4100 |

C

FIG. 54B

|  |      |
|--|------|
| MKKIENLKWKNVSFKSLEIDPDAGVVLVSVEKFSEEIEDLVRLLLEKKTRF  |      |
| RVIVNGVQKSNGLRGKILSLLNGNVPYIKDVVFEGNRLILKVLGDFARD    | 100  |
| RIASKLRSTKKQLDELLPPGTEIMLEVVEPPEDLLKKEVPQPEKREEPKG   |      |
| EELKIEDENHIFGQKPRKIVFTPSKIFEYNKKTsvkgkIFKIEKIEGKRT   | 200  |
| VLLIYLTGDGDSLICKVFNDVEKVEGKVSVDVIVATGDLLLENGEPTLY    |      |
| VKGITKLPEAKRMDKSPVKRVELHAHTKFSQDAITDVNEYVKRAKEWGF    | 300  |
| PAIALTDHGNVQAIPYFYDAAKEAGIKPIFGIEAYLVSDVEPVIRNLSDD   |      |
| STFGDATFVVLDFTTGLDPQVDEIIIEIGAVKIQGGQIVDEYHTLIKPSR   | 400  |
| EISRKSSSEITGITQEMLENKRSIEEVLPEFLGFLEDSIIIVAHNANFDYRF |      |
| LRLWIKKVMGLDWERPYIDTLALAKSLLKLRSSYSLDSVVEKLGLGPFRHH  | 500  |
| RALDDARVTAQVFLRFVEMMKIGITKLSEMEKLKDTIDYTALKPFHCTI    |      |
| LVQNKKGKLNLYKLVSDSYIKYFYGVPRILKSELINREGLLVGSACISG    | 600  |
| ELGRAALEGASDSELEEIAKFYDYIEVMPLDVIAEDEEDLDRERLKEVYR   |      |
| KLYRIAKKLNKFVVMTGdvhFLDPEDARGRAALLAPQGNRNfENQPALYL   | 700  |
| RTTEEMLEKAIEIFEDEEIAREVVNIENPNRIADMIIEVQPLEKKLHPPII  |      |
| ENADEIVRNLTMKRAYEIIYGDPLPEIVQKRVEKELNAIINHGYAVLYLIA  | 800  |
| QELVQKMSDGYVVGSRGsvGSSLVANLLGITEVNPLPPHYRCPECKYFE    |      |
| VVEDDRYGAGYDLPNKNCPRCGAPLRKDGHGIPFETFMGFEGDKVPDIDL   | 900  |
| NFSGEYQERAHRFVEELFGKDHVYRAGTINTIAERSAVGYVRSYEEKTGK   |      |
| KLRKAEMERLVSMITGVKRTTGQHPGGLMIIPKDKEYDFTPIQYPANDR    | 1000 |
| NAGVFTTHFAYETIHDDLVKIDALGHDDPTFIKMLKDLTGIDPMTIPMDD   |      |
| PDTLAIFSSVKPLGVDVPELESDVGTYGIPFEGTEFVRGMLVETRPKSFA   | 1100 |
| ELVRISGLSHGTDVWLNWARDWINLGYAKLSEVISCRDDIMNFLIHKGME   |      |
| PSLAFKIMENVRKGGITEEMESEMRRLKVPEWFIESCRIKYLFPKAHA     | 1200 |
| VAYVSMAFRIAYFKVHYPLQFYAAAYFTIKGDQFDPVLVLRGKEAIKRRLR  |      |
| ELKAMPAKDAQKKNEVSVLEVALEMILRGFSFLPPDIFKSDAKKFLIEGN   | 1300 |
| SLRIPFNKLPGLGDSVAESIIRAREEKPFTSVEDLMKRTKVKNKHIELMK   |      |
| SLGVLGDLPETEQFTLF                                    | 1367 |

FIG. 55

|   |     |
|---|-----|
| GTGCTCGCCATGATATGGAACGACACCGTTTTTTGCGTCGTAGACACAGA  |     |
| AACCACGGGAACCGATCCCTTTGCCGGAGACCGGATAGTTGAAATAGCCG  | 100 |
| CTGTTCTGTCTTCAAGGGGAAGATCTACAGAAACAAAGCGTTTCACTCT   |     |
| CTCGTGAATCCCAGAATAAGAATCCCTGCGCTGATTACAGAAAGTTCACGG | 200 |
| TATCAGCAACATGGACATCGTGGAAGCGCCAGACATGGACACAGTTTACG  |     |
| ATCTTTTCAGGGATTACGTGAAGGGAACGGTGCTCGTGTTTCACAACGCC  | 300 |
| AACTTCGACCTCACTTTTCTGGATATGATGGCAAAGGAAACGGGAACTT   |     |
| TCCAATAACGAATCCCTACATCGACACACTCGATCTTTCAGAAGAGATCT  | 400 |
| TTGGAAGGCCTCATTCTCTCAAATGGCTCTCCGAAAGACTTGGAATAAAA  |     |
| ACCACGATACGGCACCGTGCTCTTCCAGATGCCCTGGTGACCGCAAGAGT  | 500 |
| TTTTGTGAAGCTTGTGAATTTCTTGGTGAAAACAGGGTCAACGAATTCA   |     |
| TACGTGGAACCGGGG                                     | 567 |

FIG. 56

|  |     |
|--|-----|
| MLAMIWNDTVFCVVDTETTGTDPFAGDRIVEIAAVPVFKGKIYRNKAFHS |     |
| LVNPRIRIPALIQKVHGISNMDIVEAPDMDTVYDLFRDYVKGTVLVFHNA | 100 |
| NFDLTFLDMMAKETGNFPITNPYIDTLDLSEEIFGRPHSLKWLSERLGIK |     |
| TTIRHRALPDALVTARVFKLVEFLGENRVNEFIRKRG              | 189 |

FIG. 57

|  |      |
|--|------|
| GTGGAAGTTCTTTACAGGAAGTACAGGCCAAAGACTTTTTCTGAGGTTGT   |      |
| CAATCAGGATCATGTGAAGAAGGCAATAATCGGTGCTATTTCAGAAGAACA  | 100  |
| GCGTGGCCACGGATACATATTCGCCGGTCCGAGGGGAACGGGGAAGACT    |      |
| ACTCTTGCCAGAATTCTCGCAAAATCCCTGAACTGTGAGAACAGAAAGGG   | 200  |
| AGTTGAACCCTGCAATTCCTGCAGAGCCTGCAGAGAGATAGACGAGGGAA   |      |
| CCTTCATGGACGTGATAGAGCTCGACGCGGCCTCCAACAGAGGAATAGAC   | 300  |
| GAGATCAGAAGAATCAGAGACGCCGTTGGATACAGGCCGATGGAAGGTAA   |      |
| ATACAAAGTCTACATAATAGACGAAGTTCACATGCTCACGAAAGAAGCCT   | 400  |
| TCAACGCGCTCCTCAAAACACTCGAAGAACCTCCTTCCCACGTCGTGTTT   |      |
| GTGCTGGCAACGACAAACCTTGAGAAGGTTCCCTCCCACGATTATCTCGAG  | 500  |
| ATGTCAGGTTTTTCGAGTTCAGAAACATTCCCGACGAGCTCATCGAAAAGA  |      |
| GGCTCCAGGAAGTTGCGGAGGCTGAAGGAATAGAGATAGACAGGGAAGCT   | 600  |
| CTGAGCTTCATCGCAAAAAGAGCCTCTGGAGGCTTGAGAGACGCGCTCAC   |      |
| CATGCTCGAGCAGGTGTGGAAGTTCCTCGGAAGGAAAGATAGATCTCGAGA  | 700  |
| CGGTACACAGGGCGCTCGGGTTGATACCGATACAGGTTGTTTCGCGATTAC  |      |
| GTGAACGCTATCTTTTCTGGTGATGTGAAAAGGGTCTTCACCGTTCTCGA   | 800  |
| CGACGTCTATTACAGCGGGAAGGACTACGAGGTGCTCATTTCAGGAAGCAG  |      |
| TCGAGGATCTGGTCGAAGACCTGGAAAGGGAGAGAGGGGTTTACCAGGTT   | 900  |
| TCAGCGAACGATATAGTTCAGGTTTCGAGACAACTTCTGAATCTTCTGAG   |      |
| AGAGATAAAGTTTCGCCGAAGAAAAACGACTCGTCTGTAAAGTGGGTTTCGG | 1000 |
| CTTACATAGCGACGAGGTTCTCCACCACAAACGTTTCAGGAAAACGATGTC  |      |
| AGAGAAAAAACGATAATTCAAATGTACAGCAGAAAGAAGAGAAGAAAGA    | 1100 |
| AACGGTGAAGGCAAAAGAAGAAAAACAGGAAGACAGCGAGTTCGAGAAAC   |      |
| GCTTCAAAGAACTCATGGAAGAACTGAAAGAAAAGGGCGATCTCTCTATC   | 1200 |
| TTTGTGCTCTCAGCCTCTCAGAGGTGCAGTTTGACGGAGAAAAGGTGAT    |      |
| TATTTCTTTTGATTTCATCGAAAGCTATGCATTACGAGTTGATGAAGAAAA  | 1300 |
| AACTGCCTGAGCTGGAAAAACATTTTTTCTAGAAAACCTCGGGAAAAAAGTA |      |
| GAAGTTGAACTTCGACTGATGGGAAAAGAAGAAACAATCGAGAAGGTTTC   | 1400 |
| TCAGAAGATCCTGAGATTGTTTGAACAGGAGGGA                   |      |

FIG. 58

|   |     |
|---|-----|
| MEVLYRKYPKTFSEVVNQDHVKKAIIGAIQKNSVAHGFI FAGPRGTGKT  |     |
| TLARILAKSLNCENRKGVEPCNSCRACREIDEGTFMDVIELDAASNRGID  | 100 |
| EIRRIRDAVGYPMEGKYKVYIIDEVHMLTKEAFNALLKTLEPPSHVVF    |     |
| VLATTNLEKVPPTIISRCQVFEFRNIPDELIEKRLQEVAAEGIEIDREA   | 200 |
| LSFIKRASGGLRDALTMLEQVWKFSEGTKIDLETVHRALGLIPIQVVRDY  |     |
| VNAIFSGDVKRVFTVLDDVYYSKDYEVLIQEAVEDLVEDLERERGVYQV   | 300 |
| SANDIVQVSRQLLNLLREIKFAEEKRLVCKVGSAYIATRFSTTNVQENDV  |     |
| REKNDNSNVQQKEEKETVKAKEEKQEDSEFEKRFKELMEELKEKGDLSI   | 400 |
| FVALSLSEVQFDGEKVIISFDSSKAMHYELMKKKLPELENI FSRKLGKKV |     |
| EVELRLMGKEETIEKVSQKILRLFEQEG                        | 478 |

FIG. 59



|   |      |
|---|------|
| ATGAAAGTAACCGTCACGACTCTTGAATTGAAAGACAAAATAACCATCGC  |      |
| CTCAAAAGCGCTCGCAAAGAAATCCGTGAAACCCATTCTTGCTGGATTTC  | 100  |
| TTTTCGAAGTGAAAGATGGAAATTTCTACATCTGCGCGACCGATCTCGAG  |      |
| ACCGGAGTCAAAGCAACCGTGAATGCCGCTGAAATCTCCGGTGAGGCACG  | 200  |
| TTTTGTGGTACCAGGAGATGTCATTGAGAAGATGGTCAAGGTTCTCCAG   |      |
| ATGAGATAACGGAACTTTCTTTAGAGGGGGATGCTCTTGTTATAAGTTCT  | 300  |
| GGAAGCACCGTTTTTCAGGATCACCACCATGCCCGCGGACGAATTTCCAGA |      |
| GATAACGCCTGCCGAGTCTGGAATAACCTTCGAAGTTGACACTTCGCTCC  | 400  |
| TCGAGGAAATGGTTGAAAAGGTCATCTTCGCCGCTGCCAAAGACGAGTTC  |      |
| ATGCGAAATCTGAATGGAGTTTTCTGGGAACTCCACAAGAATCTTCTCAG  | 500  |
| GCTGGTTGCAAGTGATGGTTTCAGACTTGCACTTGCTGAAGAGCAGATAG  |      |
| AAAACGAGGAAGAGGCGAGTTTCTTGCTCTCTTTGAAGAGCATGAAAGAA  | 600  |
| GTTCAAAACGTGCTGGACAACACAACGGAGCCGACTATAACGGTGAGGTA  |      |
| CGATGGAAGAAGGGTTTTCTGTGTCGACAAATGATGTAGAAACGGTGATGA | 700  |
| GAGTGGTCGACGCTGAATTTCCCGATTACAAAAGGGTGATCCCCGAAACT  |      |
| TTCAAAACGAAAGTGGTGGTTTTCCAGAAAAGAACTCAGGGAATCTTTGAA | 800  |
| GAGGGTGATGGTGATTGCCAGCAAGGGAAGCGAGTCCGTGAAGTTCGAAA  |      |
| TAGAAGAAAACGTTATGAGACTTGTGAGCAAGAGCCCGGATTATGGAGAA  | 900  |
| GTGGTCGATGAAGTTGAAGTTCAAAAAGAAGGGGAAGATCTCGTGATCGC  |      |
| TTTCAACCCGAAGTTCATCGAGGACGTTTTGAAGCACATTGAGACTGAAG  | 1000 |
| AAATCGAAATGAACTTCGTTGATTCTACCAGTCCATGTCAGATAAATCCA  |      |
| CTCGATATTTCTGGATACCTTTACATAGTGATGCCCATCAGACTGGCA    | 1098 |

FIG. 60

|   |     |
|---|-----|
| MKVTVTTLLELKDKITIASKALAKKSVKPILAGFLFEVKDGNFYICATDLE |     |
| TGVKATVNAAEISGEARFVVPGDVIQKMKVLPDEITELSLEGDALVISS   | 100 |
| GSTVFRITTMPADEFPEITPAESGITFEVDTSLLEEMVEKVIFAAAKDEF  |     |
| MRNLNGVFWELHKNLLRLVASDGFRLALAEEQIENEEASFLLSLKSMKE   | 200 |
| VQNVLDNTTEPTITVRYDGRRVSLSTNDVETVMRVVDAEFPDYKRVIPET  |     |
| FKTKVVVSRKELRESLKRMVIAASKGSESVKFEIEENVMLVSKSPDYGE   | 300 |
| VVDEVEVQKEGEDLVIAFNPKFIEDVLKHIEETEEIEMNFVDSTSPCQINP |     |
| LDISGYLYIVMPIRLA                                    | 366 |

FIG. 61

|   |     |
|---|-----|
| ATGCCAGTCACGTTTCTCACAGGTACTGCAGAACTCAGAAGGAAGAATT   |     |
| GATAAAGAACTCCTGAAGGATGGTAACGTGGAGTACATAAGGATCCATC   | 100 |
| CGGAGGATCCCGACAAGATCGATTTTCATAAGGTCTTTACTCAGGACAAAG |     |
| ACGATCTTTTCCAACAAGACGATCATTGACATCGTCAATTTTCGATGAGTG | 200 |
| GAAAGCACAGGAGCAGAAGCGTCTCGTTGAACTTTTGAAAAACGTACCGG  |     |
| AAGACGTTTCATATCTTCATCCGTTCTCAAAAAACAGGTGGAAAGGGAGTA | 300 |
| GCGCTGGAGCTTCCGAAGCCATGGGAAACGGACAAGTGGCTTGAGTGGAT  |     |
| AGAAAAGCGCTTCAGGGAGAATGGTTTGCTCATCGATAAAGATGCCCTTC  | 400 |
| AGCTGTTTTTCTCCAAGGTGGAACGAACGACCTGATCATAGAAAGGGAG   |     |
| ATTGAAAAACTGAAAGCTTATTCCGAGGACAGAAAGATAACGGTAGAAGA  | 500 |
| CGTGGAAGAGGTCGTTTTTACCTATCAGACTCCGGGATACGATGATTTTT  |     |
| GCTTTGCTGTTTCCGAAGGAAAAAGGAAGCTCGCTCACTCTCTTCTGTCTG | 600 |
| CAGCTGTGGAAAACCAAGAGTCCGTGGTGATTGCCACTGTCTTGCGAA    |     |
| TCACTTCTTGATCTCTTCAAAATCCTCGTTCTTGAGACAAAGAAAAGAT   | 700 |
| ACTACACCTGGCCTGATGTGTCCAGGGTGTCCAAAGAGCTGGGAATTCCC  |     |
| GTTCCCTCGTGTGGCTCGTTTCCTCGGTTTCTCCTTTAAGACCTGGAAATT | 800 |
| CAAGGTGATGAACCACCTCCTCTACTACGATGTGAAGAAGGTTAGAAAGA  |     |
| TACTGAGGGATCTCTACGATCTGGACAGAGCCGTGAAAAGCGAAGAAGAT  | 900 |
| CCAAAACCGTTCTTCCACGAGTTCATAGAAGAGGTGGCACTGGATGTATA  |     |
| TTCTCTTCAGAGAGATGAAGAA                              | 972 |

FIG. 62

|   |     |
|---|-----|
| MPVTFLTGTAEQTKEELIKLLKDG NVEYIRIHPEDPDKIDFIRSLLRTK  |     |
| TIFSNKTIIDIVNFDEWKAQEQRLVELLKNVPEDVHIFIRSQKTGGKGV   | 100 |
| ALELPKPWETDKWLEWIEKRFRENGLLIDKDALQLFFSKVGTNDLI IERE |     |
| IEKLKAYSEDRKITVEDVEEVVFTYQTPGYDDFCFAVSEGKRKLAHSLLS  | 200 |
| QLWKTTESVVIATVLANHFLDLFKILVLVTKKRYYTWPDVSRVSKELGIP  |     |
| VPRVARFLGFSFKTWKFKVMNHLLEYDVKKVRKILRDLYDLDRAVKSEED  | 300 |
| PKPFFHEFIEEVALDVYSLQRDEE                            |     |

FIG. 63

|   |     |
|---|-----|
| ATGAACGATTTGATCAGAAAGTACGCTAAAGATCAACTGGAACTTTGAA   |     |
| AAGGATCATAGAAAAGTCTGAAGGAATATCCATCCTCATAAATGGAGAAG  | 100 |
| ATCTCTCGTATCCGAGAGAAGTATCCCTTGAACCTCCCGAGTACGTGGAG  |     |
| AAATTTCCCCCGAAGGCCTCGGATGTTCTGGAGATAGATCCCGAGGGGGA  | 200 |
| GAACATAGGCATAGACGACATCAGAACGATAAAGGACTTCCTGAACTACA  |     |
| GCCCCGAGCTCTACACGAGAAAGTACGTGATAGTCCACGACTGTGAAAGA  | 300 |
| ATGACCCAGCAGGCGGCGAACGCGTTTCTGAAGGCCCTTGAAGAACCACC  |     |
| AGAATACGCTGTGATCGTTCTGAACACTCGCCGCTGGCATTATCTACTGC  | 400 |
| CGACGATAAAGAGCCGAGTGTTTCTGAGTGGTTGTGAACGTTCCAAAGGAG |     |
| TTCAGAGATCTCGTGAAAGAGAAAATAGGAGATCTCTGGGAGGAACTTCC  | 500 |
| ACTTCTTGAGAGAGACTTCAAACCGGCTCTCGAAGCCTACAAACTTGGTG  |     |
| CGGAAAAAATTTCTGGATTGATGGAAAGTCTCAAAGTTTTGGAGACGGAA  | 600 |
| AAACTCTTGAAAAAGGTCCTTTCAAAGGCCTCGAAGGTTATCTCGCATG   |     |
| TAGGGAGCTCCTGGAGAGATTTTCAAAGGTGGAATCGAAGGAATTCTTTG  | 700 |
| CGCTTTTTGATCAGGTGACTAACACGATAACAGGAAAAGACGCGTTTCTT  |     |
| TTGATCCAGAGACTGACAAGAATCATTCTCCACGAAAACACATGGGAAAG  | 800 |
| CGTTGAAGATCAAAAAAGCGTGTCTTTCCTCGATTCAATTCTCAGGGTGA  |     |
| AGATAGCGAATCTGAACAACAACTCACTCTGATGAACATCCTCGCGATA   | 900 |
| CACAGAGAGAGAAAGAGAGGTGTCAACGCTTGGAGC                |     |

FIG. 64

|  |     |
|--|-----|
| MNDLIRKYAKDQLETLKRIIEKSEGISILINGEDLSYPREVSLELPEYVE |     |
| KFPPKASDVLEIDPEGENIGIDDIRTIKDFLNYSPELYTRKYVIVHDCER | 100 |
| MTQQAANAFLKALEEPPEYAVIVLNTRRWHYLLPTIKSRVFRVVVNPKE  |     |
| FRDLVKEKIGDLWEELPLLERDFKTALEYKLGAEKLSGLMESLKVLETE  | 200 |
| KLLKKVLSKLEGYLACRELLERFSKVESKEFFALFDQVTNTITGKDAFL  |     |
| LIQRLTRIILHENTWESVEDKSVSFLDSILRVKIANLNNKLTLMNILAIH | 300 |
| RERKRGVNAWS  |     |

FIG. 65

ATGTCCTTTCTTCAACAAGATCATACTCATAGGAAGACTCGTGAGAGATCC  
 CGAAGAGAGATACACGCTCAGCGGAACTCCAGTCACCACCTTCACCATAG 100  
 CGGTGGACAGGGTTCCCAGAAAGAACGCGCCGGACGACGCTCAAACGACT  
 GATTTCTTCAGGATCGTCACCTTTGGAAGACTGGCAGAGTTCGCTAGAAC 200  
 CTATCTCACCAAAGGAAGGCTCGTTCTCGTCGAAGGTGAAATGAGAATGA  
 GAAGATGGGAAACACCCACTGGAGAAAAGAGGGTATCTCCGGAGGTTGTC 300  
 GCAAACGTTGTTAGATTCATGGACAGAAAACCTGCTGAAACAGTTAGCGA  
 GACTGAAGAGGAGCTGGAAATACCGGAAGAAGACTTTTCCAGCGATACCT 400  
 TCAGTGAAGATGAACCACCATT

**FIG. 66**

MSFFNKIILIGRLVRDPEERYTTLSGTPVTTFTIAVDRVPRKNAPDDAQT  
 DFFRIVTFGRLAEFARTYLTGRLVLVEGEMRMRRWETPTGEKRVSPVV 100  
 ANVVRFMDRKPAETVSETEEELEIPEEDFSSDTFSEDEPPF

**FIG. 67**

|   |      |
|---|------|
| ATGCGTGTTCCTCCCGCACAACTTAGAGGCCGAAGTTGCTGTGCTCGGAAG | 100  |
| CATATTGATAGATCCGTCGGTAATAAACGACGTTCTTGAAATTTTGAGCC  |      |
| ACGAAGATTTCTATCTGAAAAACACCAACACATCTTCAGAGCGATGGAA   | 200  |
| GAGCTTTACGACGAAGGAAAACCGGTGGACGTGGTTTCCGTCTGTGACAA  |      |
| GCTTCAAAGCATGGGAAAACCTCGAGGAAGTAGGTGGAGATCTGGAAGTGG | 300  |
| CCCAGCTCGCTGAGGCTGTGCCCAGTTCTGCACACGCACCTTCACTACGCG |      |
| GAGATCGTCAAGGAAAAATCCATTCTGAGGAACTCATTGAGATCTCCAG   | 400  |
| AAAAATCTCAGAAAGTGCCTACATGGAAGAAGATGTGGAGATCCTGCTCG  |      |
| ACAACGCAGAAAAGATGATCTTCGAGATCTCAGAGATGAAAACGACAAAA  | 500  |
| TCCTACGATCATCTGAGAGGCATCATGCACCGGGTGTGTTGAAAACCTGGA |      |
| GAACTTCAGGGAAAGAGCCAACCTTATAGAACCCGGTGTGCTCATAACGG  | 600  |
| GACTACCAACGGGATTCAAAAGTCTGGACAAACAGACCACAGGGTTCAC   |      |
| AGCTCCGATCTGGTGATAATAGCAGCGAGACCCTCCATGGGAAAAACCTC  | 700  |
| CTTCGCACTCTCAATAGCGAGGAACATGGCTGTCAATTTCGAAATCCCCG  |      |
| TCGGAATATTCACTCTCGAGATGTCCAAGGAACAGCTCGCTCAAAGACTA  | 800  |
| CTCAGCATGGAGTCCGGTGTGGATCTTTACAGCATCAGAACAGGATACCT  |      |
| GGATCAGGAGAAGTGGGAAAGACTCACAATAGCGGCTTCTAAACTCTACA  | 900  |
| AAGCACCCTAGTTGTGGACGATGAGTCACTCCTCGATCCGCGATCGTTG   |      |
| AGGGCAAAGCGAGAAGGATGAAAAAGAATACGATGTAAAAGCCATTTT    | 1000 |
| TGTCGACTATCTCCAGCTCATGCACCTGAAAGGAAGAAAAGAAAGCAGAC  |      |
| AGCAGGAGATATCCGAGATCTCGAGATCTCTGAAGCTCCTTGCGAGGGAA  | 1100 |
| CTCGACATAGTGGTGATAGCGCTTTTACAGCTTTTCGAGGGCCGTAGAACA |      |
| GAGAGAAGACAAAAGACCGAGGCTGAGTGACCTCAGGGAATCCGGTGCGA  | 1200 |
| TAGAACAGGACGCAGACACAGTCATCTTCATCTACAGGGAGGAATATTAC  |      |
| AGGAGCAAAAAATCCAAAGAGGAAAGCAAGCTTCACGAACCTCACGAAGC  | 1300 |
| TGAAATCATAATAGGTAAACAGAGAAACGGTCCCGTTGGAACGATCACTC  |      |
| TGATCTTCGACCCAGAACGGTTACGTTCCATGAAGTCGATGTGGTGCAT   | 1353 |
| TCA   |      |

FIG. 68

|   |     |
|---|-----|
| MRVPPHNLEAEVAVLGSILIDPSVINDVLEILSHEDFYLLKKHQHIFRAME | 100 |
| ELYDEGKPDVSVCDKLQSMGKLEEVGGDLEVAQLAEAVPSSAHALHYA    |     |
| EIVKEKSILRKLIEISRKISESAYMEEDVEILLDNAEKMIFEISEMKTTK  | 200 |
| SYDHLRGIMHRVFENLENFRERANLIEPGVLITGLPTGFKSLDKQTTGFH  |     |
| SSDLVIAARPSMGKTSFALSIARNMAVNFEIPVGIFSLEMSKEQLAQR    | 300 |
| LSMESGVDLYSIRTGYLDQEKWERLTIAASKLYKAPIVVDDESLLDPRSL  |     |
| RAKARRMKKEYDVKAIFVDYLQMLHLKGRKESRQOEISEISRSLKLLARE  | 400 |
| LDIVVIALSQLSRAVEQREDKRPRLSDLRESGAIEQDADTVIFIYREEYY  |     |
| RSKKSKEESKLHEPHEAEIIIGKQRNGPVGITITLIFDPRTVTFHEVDVVH | 451 |
| S   |     |

FIG. 69

|   |      |
|---|------|
| GTGATTCTCGAGAGGTCATCGAGGAAATAAAAGAAAAGGTTGACATCGT     |      |
| AGAGGTCATTTCCGAGTACGTGAATCTTACCCGGGTAGGTTCTCTCTACA    | 100  |
| GGGCTCTCTGTCCCTTTTCATTTCAGAAACCAATCCTTCTTTCTACGTTTCAT |      |
| CCGGGTTTGAAGATATACCATTTGTTTCGGCTGCGGTGCGAGTGGAGACGT   | 200  |
| CATCAAATTTCTTCAAGAAATGGAAGGGATCAGTTTCCAGGAAGCGCTGG    |      |
| AAAGACTTGCCAAAAGAGCTGGGATTGATCTTTCTCTCTACAGAACAGAA    | 300  |
| GGGACTTCTGAATACGGAAAATACATTCGTTTGTACGAAGAAACGTGGAA    |      |
| AAGGTACGTCAAAGAGCTGGAGAAATCGAAAGAGGCCAAAAGACTATTTAA   | 400  |
| AAAGCAGAGGCTTCTCTGAAGAAGATATAGCAAAGTTCGGCTTTGGGTAC    |      |
| GTCCCCAAGAGATCCAGCATCTCTATAGAAGTTGCAGAAGGCATGAACAT    | 500  |
| AACACTGGAAGAAGTTGTTCAGATACGGTATCGCGCTGAAAAAGGGTGATC   |      |
| GATTCGTTGATAGATTCTGAAGGAAGAATCGTTGTTCCAATAAAGAACGAC   | 600  |
| AGTGGTCATATTGTGGCTTTTGGTGGGCGTGCTCTCGGCAACGAAGAACC    |      |
| GAAGTATTTGAACTCTCCAGAGACCAGGTATTTTTCGAAGAAGAAGACCC    | 700  |
| TTTTTCTCTTCGATGAGGCGAAAAAGTGGCAAAAGAGGTTGGTTTTTTC     |      |
| GTCATCACCGAAGGCTACTTCGACGCGCTCGCATTTCAGAAAGGATGGAAT   | 800  |
| ACCAACGGCGGTCGCTGTTCTTGGGGCGAGTCTTTCAGAGAGGCGATTCT    |      |
| TAAACTTTTCGGCGTATTCGAAAAACGTCATACTGTGTTTTCGATAATGAC   | 900  |
| AAAGCAGGCTTCAGAGCCACTCTCAAATCCCTCGAGGATCTCCTAGACTA    |      |
| CGAATTCAACGTGCTTGTGGCAACCCCTCTCCTTACAAAGACCCAGATG     | 1000 |
| AACTCTTTCAGAAAGAAGGAGAAGGTTTCATTGAAAAAGATGCTGAAAAAC   |      |
| TCGCGTTCGTTTCGAATATTTTCTGGTGACGGCTGGTGAGGTCTTCTTTGA   | 1100 |
| CAGGAACAGCCCCGCGGTGTGAGATCCTACCTTCTTTCCTCAAAGGTT      |      |
| GGGTCCAAAAGATGAGAAGGAAAGGATATTTGAAACACATAGAAAATCTC    | 1200 |
| GTGAATGAGGTTTCATCTTCTCTCCAGATACCAGAAAACCAGATTTTGAA    |      |
| CTTTTTTGAAGCGACAGGTCTAACACTATGCCTGTTTCATGAGACCAAGT    | 1300 |
| CGTCAAAGGTTTACGATGAGGGGAGAGGACTGGCTTATTTGTTTTTGAAC    |      |
| TACGAGGATTTGAGGGAAAAGATTCTGGAAGTGGACTTAGAGGTACTGGA    | 1400 |
| AGATAAAAACGCGAGGGAGTTTTTCAAGAGAGTCTCACTGGGAGAAGATT    |      |
| TGAACAAAGTCATAGAAAACCTCCCAAAGAGCTGAAAGACTGGATTTTTT    | 1500 |
| GAGACAATAGAAAGCATTCTCTCTCCAAAGGATCCCGAGAAATTCCTCGG    |      |
| TGACCTCTCCGAAAAGTTGAAAATCCGACGGATAGAGAGACGTATCGCAG    | 1600 |
| AAATAGATGATATGATAAAGAAAGCTTCAAACGATGAAGAAAGGCGTCTT    |      |
| CTTCTCTCTATGAAAGTGGATCTCTCAGAAAAATAAAGAGGAGG          | 1695 |

FIG. 70

|   |     |
|---|-----|
| MIPREVIEEIKEKVDIVEVISEYVNLTRVGSSYRALCPFHSETNPSFYVH  |     |
| PGLKIYHCFGCGASGDVIKFLQEMEGISFQEALERLAKRAGIDLSLYRTE  | 100 |
| GTSEYGKYIRLYEETWKRYVKELEKSKEAKDYLSRGSFSEEDIAKFGFGY  |     |
| VPKRSSISIEVAEGMNITLEELVRYGIALKKGDRFVDRFEGRIVVPIKND  | 200 |
| SGHIVAFGGRALGNEEPKYLNSPETRYFSKKKTLFLFDEAKKVAKEVGFF  |     |
| VITEGYFDALAFRKDGIPAVAVLGASLSREAILKLSAYSKNVILCFDND   | 300 |
| KAGFRATLKSLEDLLDYEFNVLVATPSPYKDPDEFQKEGEGSLKKMLKN   |     |
| SRSFEYFLVTAGEVFFDRNSPAGVRSYLSFLKGWVQKMRRKGYLKHIEENL | 400 |
| VNEVSSSLQIPENQILNFFESDRSNTMPVHETKSSKVYDEGRGLAYLFLN  |     |
| YEDLREKILELDLEVLEDKNAREFFKRVSLGEDLNKVIENFPKELKDWIF  | 500 |
| ETIESIPPPKDPEKFLGDLSEKLKIRRIERRIAEIDDMIKKASNDEERRL  |     |
| LLSMKVDLLRKIKRR                                     | 565 |

FIG. 71

|  |     |
|--|-----|
| ATGGCTCTACACCCGGCTCACCTGGGGCAATAATCGGGCAGGAGCCGT   |     |
| TCTCGCCCTCCTTCCCCGCCTCACCGCCAGACCCTGCTCTTCTCCGGCC  | 100 |
| CCGAGGGGGTGGGGCGGCGCACCGTGGCCCGCTGGTACGCCTGGGGGCTC |     |
| AACCGCGGCTTCCCCCGCCCTCCCTGGGGGAGCACCCGGACGTCCTCGA  | 200 |
| GGTGGGGCCCAAGGCCCGGGACCTCCGGGGCCGGGCCGAGGTGCGGCTGG |     |
| AGGAGGTGGCGCCCTCTTGGAGTGGTGCTCCAGCCACCCCCGGGAGCGG  | 300 |
| GTGAAGGTGGCCATCCTGGACTCGGCCACCTCCTCACCGAGGCCGCCGC  |     |
| CAACGCCCTCCTCAAGCTCCTGGAGGAGCCCCCTTCTACGCCCGCATCG  | 400 |
| TCCTCATCGCCCCAAGCCGCGCCACCCTCCTCCCCACCCTGGCCTCCCGG |     |
| GCCACGGAGGTGGCATTGCCCCCGTGCCCGAGGAGGCCCTGCGCGCCCT  | 500 |
| CACCCAGGACCCGGAGCTCCTCCGCTACGCCGCGGGGGCCCCGGGCCGCC |     |
| TCCTTAGGGCCCTCCAGGACCCGGAGGGGTACCGGGCCCGCATGGCCAGG | 600 |
| GCGCAAAGGGTCTGAAAGCCCCGCCCTGGAGCGCCTCGCTTTGCTTCG   |     |
| GGAGCTTTTGGCCGAGGAGGAGGGGGTCCACGCCCTCCACGCCGTCTTAA | 700 |
| AGCGCCCGGAGCACCTCCTTGCCCTGGAGCGGGCGCGGGAGGCCCTGGAG |     |
| GGGTACGTGAGCCCCGAGCTGGTCCTCGCCCGGCTGGCCTTAGACTTAGA | 800 |
| GACA   |     |

FIG. 72

|  |     |
|--|-----|
| MALHPAHPGAIIGHEAVLALLPRLTAQTLFSGPEGVGRRTVARWYAWGL  |     |
| NRGFPPPSLGEHPDVLEVGPKARDLRGRAEVRLEEVAPLLWCSSHPRER  | 100 |
| VKVAILDSAHLLTEAAANALLKLLEEPSYARIVLIAPSRATLLPTLASR  |     |
| ATEVAFAPVPEEALRALTDPELLRYAAGAPGRLLRALQDPEGYRARMAR  | 200 |
| AQRVLKAPPLERLALLRELLAEEEGVHALHAVLKRPEHLLALERAREALE |     |
| GYVSPELVLARLALDLET                                 | 268 |

FIG. 73

|   |     |
|---|-----|
| ATGCTGGACCTGAGGGAGGTGGGGGAGGCGGAGTGGAAGGCCCTAAAGCC  |     |
| CCTTTTGGAAAGCGTGCCCGAGGGCGTCCCGTCCTCCTCCTGGACCCTA   | 100 |
| AGCCAAGCCCCCTCCCGGGCGGCCTTCTACCGGAACCGGGAAAGGCGGGAC |     |
| TTCCCCACCCCCAAGGGGAAGGACCTGGTGCGGCACCTGGAAAACCGGGC  | 200 |
| CAAGCGCCTGGGGCTCAGGCTCCCGGGCGGGGTGGCCAGTACCTGGCCT   |     |
| CCCTGGAGGGGGACCTCGAGGCCCTGGAGCGGGAGCTGGAGAAGCTTGCC  | 300 |
| CTCCTCTCCCCACCCCTCACCTGGAGAAGGTGGAGAAGGTGGTGGCCCT   |     |
| GAGGCCCCCCCTCACGGGCTTTGACCTGGTGCCTCCGTCTCTGGAGAAGG  | 400 |
| ACCCCAAGGAGGCCCTCCTGCGCCTAGGCGGCCTCAAGGAGGAGGGGGAG  |     |
| GAGCCCCTCAGGCTCCTCGGGGCCCTCTCCTGGCAGTTGCCCCCTCTCGC  | 500 |
| CCGGGCCTTCTTCTCCTCCGGGAAAACCCAGGCCCAAGGAGGAGGACC    |     |
| TCGCCCCGCTCGAGGCCACCCCTACGCCGCCCGCCGCGCCCTGGAGGCG   | 600 |
| GCGAAGCGCCTCACGGAAGAGGCCCTCAAGGAGGCCCTGGACGCCCTCAT  |     |
| GGAGGCGGAAAAGAGGGGCAAGGGGGGAAAGACCCGTGGCTCGCCCTGG   | 700 |
| AGGCGGCGGTCCTCCGCCTCGCCCGTTGA                       |     |

**FIG. 74**

|   |     |
|---|-----|
| MVIAFTGDPFLAREALLEEARLRGLSRFTEPTPEALAQALAPGLFGGGGA  |     |
| MLDLREVGEAEWKALKPLLESVPEGVPVLLLDPKPSPSRAAFYRNRERRD  | 100 |
| FPTPKGKDLVRHLENRAKRLGLRLPGGVAQYLASLEGDLEALERELEKLA  |     |
| LLSPPLTLEKVEKVVALRPPLTGFDLVRVLEKDPKEALLRLGGLKEEGE   | 200 |
| EPLRLLGALSQWQFALLARAFFLLRENPRPKEEDLARLEAHPYAARRALEA |     |
| AKRLTEEALKEALDALMEAEKRAKGKDPWLALEAAVLRRLAR          | 292 |

**FIG. 75**



|   |     |
|---|-----|
| ATGGCTCGAGGCCTGAACCGCGTTTTCTCATCGGCGCCCTCGCCACCCG   |     |
| GCCGGACATGCGCTACACCCGGCGGGGCTCGCCATTTTGGACCTGACCC   | 100 |
| TCGCCGGTCAGGACCTGCTTCTTTCCGATAACGGGGGGGAACCGGAGGTG  |     |
| TCCTGGTACCACCGGGTGAGGCTCTTAGGCCGCCAGGCGGAGATGTGGGG  | 200 |
| CGACCTCTTGGAACCAAGGGCAGCTCGTCTTCGTGGAGGGCCGCCTGGAGT |     |
| ACCGCCAGTGGGAAGGGAGGGGGAGAAGCGGAGCGAGCTCCAGATCCGG   | 300 |
| GCCGACTTCCGGACCCCCTGGACGACCGGGGAAGAAGCGGGCGGAGGAC   |     |
| AGCCGGGGCCAGCCCAGGCTCCGCGCCGCCCTGAACCAGGTCTTCCTCAT  | 400 |
| GGGCAACCTGACCCGGGACCCGGAACCTCCGCTACACCCCCAGGGCACCG  |     |
| CGGTGGCCCGGCTGGGCCTGGCGGTGAACGAGCGCCGCAGGGGGCGGAG   | 500 |
| GAGCGCACCCACTTCGTGGAGGTTTCAGGCCTGGCGCGACCTGGCGGAGTG |     |
| GGCCGCCGAGCTGAGGAAGGGCGACGGCCTTTTCGTGATCGGCAGGTTGG  | 600 |
| TGAACGACTCCTGGACCAGCTCCAGCGGCGAGCGGCGCTTCCAGACCCGT  |     |
| GTGGAGGCCCTCAGGCTGGAGCGCCCCACCCGTGGACCTGCCCAGGCCTG  | 700 |
| CCCAGGCCGGCGGAACAGGTCCCGCGAAGTCCAGACGGGTGGGGTGGACA  |     |
| TTGACGAAGGCTTGGAAGACTTTCCGCCGGAGGAGGATTTGCCGTTTTGA  | 800 |
| GCACGAA   |     |

FIG. 76

|  |     |
|--|-----|
| MARGLNRVFLIGALATRPDMRYTPAGLAILDLTLAGQDLLSDNGGEPEV    |     |
| SWYHRVRLGRQAEMWGDLLDQGQLVFVEGRLEYRQWEREGEKRSELQIR    | 100 |
| ADFLDPLDDRGGKKRAEDSRGQPRRLRAALNQVFLMGNLTRDPELRYTPQGT |     |
| AVARLGLAVNERRQGAERTHFVEVQAWRDLAEWAAELRKGDGLFVIGRL    | 200 |
| VNDSWTSSSGERRFQTRVEALRLERPTRGPAQACPGRNRNRSREVQTGGVD  |     |
| IDEGLEDFPPEEDLPF                                     | 266 |

FIG. 77

|  |     |
|--|-----|
| AATTCCGACATTTCAATTGAATCGTTTATTCCGCTTGAAAAAGAAGGCAA   | 100 |
| GTTGCTCGTTGATGTGAAAAGACCGGGGAGCATCGTACTGCAGGCGCGCT   |     |
| TTTTCTCTGAAATCGTGAAAAAACTGCCGCAACAAACGGTGGAAATCGAA   | 200 |
| ACGGAAGACAACCTTTTTGACGATCATCCGCTCGGGGCACTCAGAATTCCG  |     |
| CCTCAATGGGCTAAACGCCGACGAATATCCGCGCCTGCCGCAAATTGAAG   | 300 |
| AAGAAAACGTGTTTCAAATCCCGGCTGATTTATTGAAAACCGTGATTTCGG  |     |
| CAAACGGTGTTTCGCCGTTTCTACATCGGAAACGCGCCCAATCTTGACAGG  | 400 |
| TGTCAACTGGAAAGTTGAACATGGCGAGCTTGTCTGCACAGCGACCGACA   |     |
| GTCATCGCTTAGCCATGCGCAAAGTGAAAATTGAGTCGGAAAATGAAGTA   | 500 |
| TCATACAACGTCGTCATCCCTGGAAAAAGTCTTAATGAGCTCAGCAAAAT   |     |
| TTTGGATGACGGCAACCACCCGGTGGACATCGTCATGACAGCCAATCAAG   | 600 |
| TGCTATTTAAGGCCGAGCACCTTCTCTTCTTTTCCCGGCTGCTTGACGGC   |     |
| AACTATCCGGAGACGGCCCGCTTGATTCCAACAGAAAGCAAAACGACCAT   | 700 |
| GATCGTCAATGCAAAAGAGTTTCTGCAGGCAATCGACCGAGCGTCCTTGC   |     |
| TTGCTCGAGAAGGAAGGAACAACGTTGTGAAACTGACGACGCTTCCTGGA   | 800 |
| GGAATGCTCGAAATTTCTTCGATTTCTCCGAGATCGGGAAAGTGACGGAG   |     |
| CAGCTGCAAACGGAGTCTCTTGAAGGGGAAGAGTTGAACATTTTCGTTTCAG | 900 |
| CGCGAAATATATGATGGACGCGTTGCGGGCGCTTGATGGAACAGACATTT   |     |
| CAAATCAGCTTCACTGGGGCCATGCGGCCGTTCTGTGTCGCCCGCTTCA    | 992 |
| ACCGATTGATGCTTCAGCTCATTTTGCCGGTGAGAACATAT            |     |

FIG. 78

|  |     |
|--|-----|
| NSDISIIIESFIPLEKEGKLLVDVKRPGSIVLQARFFSEIVKKLPQQTVEI  | 100 |
| ETEDNFLTIIIRSGHSEFRLNGLNADEYPRLPQIEEENVFQIPADLLKTVI  |     |
| RQTVFAVSTSETRPILTGVDNWKVEHGELVCTATDSHRLAMRKVKII ESEN | 200 |
| EVSYNVVIPGKSLNELSKIILDDGNHPVDIVMTANQVLFKAEHLLFFSRL   |     |
| LDGNYPETARLIPTESKTTMIVNAKEFLOAIDRASLLAREGRNNVVKLTT   | 300 |
| LPGGMLEISSISPEIGKVTEQLQTESLEGEELNISFSAKYMMDALRALDG   |     |
| TDIQISFTGAMRPFLRLPLHTDSMLQLILPVRTY                   |     |

FIG. 79

|  |     |
|--|-----|
| ATGATTAACCGCGTCATTTTGGTCGGCAGGTTAACGAGAGATCCGGAGTT   |     |
| GCGTTACACTCCAAGCGGAGTGGCTGTTGCCACGTTTACGCTCGCGGTCA   | 100 |
| ACCGTCCGTTTACAAATCAGCAGGGCGAGCGGGAAACGGATTTTATTCAA   |     |
| TGTGTCGTTTGGCGCCGCCAGGCGGAAAACGTCGCCAACTTTTTGAAAAA   | 200 |
| GGGGAGCTTGGCTGGTGTGTCGATGGCCGACTGCAAACCCGCAGCTATGAAA |     |
| ATCAAGAAGGTCGGCGTGTGTACGTGACGGAAGTGGTGGCTGATAGCGTC   | 300 |
| CAATTTCTTGAGCCGAAAGGAACGAGCGAGCAGCGAGGGGCGACAGCAGG   |     |
| CGGCTACTATGGGGATCCATTCCCATTGCGGCAAGATCAGAACCACCAAT   | 400 |
| ATCCGAACGAAAAAGGGTTTGGCCGCATCGATGACGATCCTTTTCGCCAAT  |     |
| GACGGCCAGCCGATCGATATTTCTGATGATGATTTGCCGTTT           | 492 |

**FIG. 80**

|  |     |
|--|-----|
| MINRVILVGRLTRDPELRYTPSGVAVATFTLAVNRPFTNQSYENQEGRRV |     |
| YVTEVVADSVQFLEPKGTSEQRGATAGGYQGERETDFIQCVVWRRQAEN  | 100 |
| VANFLKKGSLAGVDGRLQTRGDPFPFGQDQNHQYPNEKGFGRIDDDPFAN |     |
| DGQPIDISDDDLPF                                     | 164 |

**FIG. 81**

|   |      |
|---|------|
| ATGCTGGAACGCGTATGGGGAAACATTGAAAAACGGCGTTTTTCTCCCCT  |      |
| TTATTTATTATACGGCAATGAGCCGTTTTTATTAACGGAAACGTATGAGC  | 100  |
| GATTGGTGAACGCAGCGCTTGGCCCCGAGGAGCGGGAGTGGAACCTTGGCT |      |
| GTGTACGACTGCGAGGAAACGCCGATCGAGGCGGCGCTTGAGGAGGCCGA  | 200  |
| GACGGTGCCGTTTTTTCGGCGAGCGGCGTGTCATTCTCATCAAGCATCCAT |      |
| ATTTTTTTTACGTCTGAAAAAGAGAAGGAGATCGAACATGATTTGGCGAAG | 300  |
| CTGGAGGCGTACTTGAAGGCGCCGTCGCCGTTTTTCGATCGTCGTCTTTTT |      |
| CGCGCCGTACGAGAAGCTTGATGAGCGAAAAAAATTACGAAGCTCGCCA   | 400  |
| AAGAGCAAAGCGAAGTCGTCATCGCCGCCCCGCTCGCCGAAGCGGAGCTG  |      |
| CGTGCTGGGTGCGGCGCCGCATCGAGAGCCAAGGGGCGCAAGCAAGCGA   | 500  |
| CGAGGCGATTGATGTCCTGTTGCGGCGGGCCGGGACGCAGCTTTCGCCT   |      |
| TGGCGAATGAAATCGATAAATTGGCCCTGTTTGCCGGATCGGGCGGAACC  | 600  |
| ATCGAGGCGGCGGCGGTTGAGCGGCTTGTCGCCCGCACGCCGGAAGAAAA  |      |
| CGTATTTGTGCTTGTCGAGCAAGTGGCGAAGCGCGACATTCCAGCAGCGT  | 700  |
| TGCAGACGTTTTTATGATCTGCTTGAAAACAATGAAGAGCCGATCAAATT  |      |
| TTGGCGTTGCTCGCCGCCCATTTCCGCTTGCTTTCGCAAGTGAAATGGCT  | 800  |
| TGCCTCCTTAGGCTACGGACAGGCGCAAATTGCTGCGGCGCTCAAGGTGC  |      |
| ACCCGTTCCGCGTCAAGCTCGCTCTTGCTCAAGCGGCCCGCTTCGCTGAC  | 900  |
| GGAGAGCTTGCTGAGGCGATCAACGAGCTCGCTGACGCCGATTACGAAGT  |      |
| GAAAAGCGGGGCGGTGATCGCCGGTTGGCCGTTGAGCTGCTTCTGATGC   | 1000 |
| GCTGGGGCGCCCGCCCGGCGCAAGCGGGGCGCCACGGCCGGCGG        |      |

**FIG. 82**

|  |     |
|--|-----|
| MLERVWGNIEKRRFSPLYLLYGNEPFLLTETYERLVNAALGPEEREWNLA |     |
| VYDCEETPIEAALEEAETVPFFGERRVILIKHPYFFTSEKEKEIEHDLAK | 100 |
| LEAYLKAPSPFSIVVFFAPYEKLDERKKITKLAKESQEVVIAAPLAEEL  |     |
| RAWVRRRIESQGAQASDEAIDVLLRRAGTQLSALANEIDKLALFAGSGGT | 200 |
| IEAAVERLVARTPEENVFVLVEQVAKRDIPALQTFYDLLENNEEPIKI   |     |
| LALLAAHFRLLSQVKWLASLGYGQAQIAAALKVHPFRVKLALAQAARFAD | 300 |
| GELAEAINELADADYEVKSGAVDRRLAVELLMLRWGARPAQAGRHGR    |     |

**FIG. 83**

|   |     |
|---|-----|
| ATGCGATGGGAACAGCTAGCGAAACGCCAGCCGGTGGTGGCGAAAATGCT  |     |
| GCAAAGCGGCTTGGA AAAAGGGCGGATTTCTCATGCGTACTTGTTTGAGG | 100 |
| GGCAGCGGGGACGGGCAAAAAGCGGCCAGTTTGTTGTTGGCGAAACGT    |     |
| TTGTTTTGTCTGTCCCCAATCGGAGTTTCCCCGTGTCTAGAGTGCCGCAA  | 200 |
| CTGCCGGCGCATCGACTCCGGCAACCACCCTGACGTCCGGGTGATCGGCC  |     |
| CAGATGGAGGATCAATCAAAAAGGAACAAATCGAATGGCTGCAGCAAGAG  | 300 |
| TTCTCGAAAACAGCGGTCGAGTCGGATAAAAAAATGTACATCGTTGAGCA  |     |
| CGCCGATCAAATGACGACAAGCGCTGCCAACAGCCTTCTGAAATTTTGG   | 400 |
| AAGAGCCGCATCCGGGGACGGTGGCGGTATTGCTGACTGAGCAATACCAC  |     |
| CGCTGCTAGGGACGATCGTTTCCCGCTGTCAAGTGCTTTCGTTCCGGCC   | 500 |
| GTTGCCCGCCGGCAGAGCTCGCCCAGGGACTTGTCGAGGAGCACGTGCCGT |     |
| TGCCGTTGGCGCTGTTGGCTGCCCATTTGACAAACAGCTTCGAGGAAGCA  | 600 |
| CTGGCGCTTGCCAAAGATAGTTGGTTTGCCGAGGCGCGAACATTAGTGCT  |     |
| ACAAATGGTATGAGATGCTGGGCAAGCCGGAGCTGCAGCTTTTGTTTTTCA | 700 |
| TCCACGACCGCTTGTTTCCGCATTTTTTGAAAGCCATCAGCTTGACCTT   |     |
| GGACTTG   | 757 |

**FIG. 84**

|   |     |
|---|-----|
| MRWEQLAKRQPVVAKMLQSGLEKGRISHAYLFEGQRTGKKAASLLAKR    |     |
| LFCLSPIGVSPCLECRNCRRIDSGNHPDVRVIGPDGGS IKKEQIEWLQQE | 100 |
| FSKTAVESDKKMYIVEHADQMTTSAANSLLKFLEEPHPGTVAVLLTEQYH  |     |
| RLLGTIVSRCQVLSFRPLPPAELAQGLVEEHVPLPLALLAAHLTNSFEEA  | 200 |
| LALAKDSWFAEARTLVLQWYEMLGKPELQLLFFIHDRLPHPFLESHQLDL  |     |
| GL  | 252 |

**FIG. 85**

|   |      |
|---|------|
| GTGGCATACCAAGCGTTATATCGCGTGTTTCGGCCGCAGCGCTTTGCGGA  |      |
| CATGGTTCGGCCAAGAACACGTGACCAAGACGTTGCAAAGCGCCCTGCTTC | 100  |
| AACATAAAATATCGCACGCTTACTTATTTTCCGGCCCGCGCGGTACAGGA  |      |
| AAAACGAGCGCAGCGAAAATTTTCGCCAAGGCGGTCAACTGTGAACAGGC  | 200  |
| GCCAGCGGCGGAGCCATGCAATGAGTGTCCAGCTTGCCTCGGCATTACGA  |      |
| ATGGAACGGTTCCTGATGTGCTGGAAATTGACGCTGCTTCCAACAACCGC  | 300  |
| GTCGATGAAATTTCGTGATATCCGTGAGAAGGTGAAATTTGCGCCAACGTC |      |
| GGCCCGCTACAAAGTGATATCATCGACGAGGTGCATATGCTGTGATCG    | 400  |
| GTGCGTTTAACGCGCTGTTGAAAACGTTGGAGGAGCCGCCGAAACACGTC  |      |
| ATTTTCATTTTGGCCACGACCGAGCCGCACAAAATTCCGGCGACGATCAT  | 500  |
| TTCCCGCTGCCAACGGTTCGATTTTCGCCGCATCCCGCTTCAGGCGATCG  |      |
| TTTCACGGCTAAAGTACGTGCAAGCGCCCAAGGTGTGAGGCGTCAGAT    | 600  |
| GAGGCATTGTCCGCCATCGCCCGTGCTGCAGACGGGGGGATGCGCGATGC  |      |
| GCTCAGCTTGCTTGATCAAGCCATTTGTTTTCAGCGACGGGAACTTCGGC  | 700  |
| TCGACGACGTGCTGGCGATGACCGGGGCTGCATCATTTGCCGCCTTATCG  |      |
| AGCTTCATCGAAGCCATCCACCGCAAAGATACAGCGGCGGTTCTTCAGCA  | 800  |
| CTTGGAACGATGATGGCGCAAGGGAAGATCCGCATCGTTTGTTGGAAG    |      |
| ACTTGATTTTGTACTATCGCGATTTATTGCTGTACAAAACCGCTCCCTAT  | 900  |
| GTGGAGGGAGCGATTCAAATTGCTGTGCTTGACGAAGCGTTCACTTCACT  |      |
| GTCGGAAATGATTCCGGTTTCCAATTTATACGAGGCCATCGAGTTGCTGA  | 1000 |
| ACAAAAGCCAGCAAGAGATGAAGTGGACAAACCACCCGCGCCTTCTGTTG  |      |
| GAAGTGGCGCTTGTGAACTTTGCCATCCATCAGCCGCCGCCCGTCGCT    | 1100 |
| GTCGGCTTCCGAGTTGGAACCGTTGATAAAGCGGATTGAAACGCTGGAGG  |      |
| CGGAATTGCGGCGCCTGAAGGAACAACCGCTGCCCTCCGTCGACCGCC    | 1200 |
| GCGCCGGTGAAAAAACTGTCCAAACCGATGAAAACGGGGGGATATAAAGC  |      |
| CCCGGTTGGCCGCATTTACGAGCTGTTGAAACAGGCGACGCATGAAGATT  | 1300 |
| TAGCTTTGGTGAAAGGATGCTGGGCGGATGTGCTCGACACGTTGAAACGG  |      |
| CAGCATAAAGTGTGCGACGCTGCCTTGCTGCAAGAGAGCGAGCCGGTTGC  | 1400 |
| AGCGAGCGCCTCAGCGTTTGTATTAAAATTCAAATACGAAATCCACTGCA  |      |
| AAATGGCGACCGATCCACAAGTTCGGTCAAAGAAAACGTGGAAGCGATT   | 1500 |
| TTGTTTGAGCTGACAAACCGCCGCTTTGAAATGGTAGCCATTCCGGAGGG  |      |
| AGAATGGGGAAAAATAAGAGAAGAGTTTCATCCGCAATAAGGACGCCATGG | 1600 |
| TGGAAAAAAGCGAAGAAGATCCGTTAATCGCCGAAGCGAAGCGGCTGTTT  |      |
| GGCGAAGAGCTGATCGAAATTAAAGAA                         | 1677 |

FIG. 86

|  |     |
|--|-----|
| VAYQALYRVFRPQRFADMVGQEHVTKTLQSALLQHKISHAYLFSGPRGTG   |     |
| KTSAAKIFAKAVNCEQAPAAEPCNECPACLGITNGTVPDVLEIDAASNNR   | 100 |
| VDEIRDIREKVKFAPTSARYKVYIIDEVHMLSIGAFNALLKTLEPPKHV    |     |
| IFILATTEPHKIPATIIISRCQRFDFRRIPLQAIIVSRLKYVASAQGVEASD | 200 |
| EALSAIARAADGGMRDALSLLDQAISFSDGKLRLDDVLAMTGAASFAALS   |     |
| SFIEAIHRKDTAAVLQHLETMMAQGKDPHRLVEDLILYYRDLLLYKTAPY   | 300 |
| VEGAIQIAVVDEAFTSLSEMI PVSNLYEAI ELLNKSQQEMKWTNHPRLLL |     |
| EVALVKLCHPSAAAPSL SASELEPLIKRIETLEAELRRLKEQPPAPPSTA  | 400 |
| APVKKLSKPMKTGGYKAPVGRIYELLKQATHEDLALVKG CWADVLDTLKR  |     |
| QHKVSHAALLQESEPVAASASAFVLKFY EIHCKMATDPTSSVKENVEAI   | 500 |
| LFELTNRRFEMVAIPEG EWGKIREEFIRNKDAMVEKSEEDPLIAEAKRLF  |     |
| GEELIEIKE  | 559 |

**FIG. 87**

|   |      |
|---|------|
| ATGGTGACAAAAGAGCAAAAAGAGCGGTTTCTCATCCTGCTTGAGCAGCT  |      |
| GAAGATGACGTCGGACGAATGGATGCCGCATTTTCGTGAGGCAGCCATTC  | 100  |
| GCAAAGTCGTGATCGATAAAGAGGAGAAAAAGCTGGCATTTTTATTTTCAG |      |
| TTCGACAACGTGCTGCCGGTTCATGTATACAAAACGTTTGCCGATCGGCT  | 200  |
| GCAGACGGCGTTCCGCCATATCGCCGCCGTCCGCCATACGATGGAGGTCG  |      |
| AAGCGCCGCGCGTAAC TGAGGCGGATGTGCAGGCGTATTGGCCGCTTTGC | 300  |
| CTTGCCGAGCTGCAAGAAGGCATGTCGCCGCTTGTCGATTGGCTCAGCCG  |      |
| GCAGACGCCTGAGCTGAAAGGAAACAAGCTGCTTGTCGTTGCCCGCCATG  | 400  |
| AAGCGGAAGCGCTGGCGATCAAACGGCGGTTCCGCCAAAAAATCGCTGAT  |      |
| GTGTACGCTTCGTTTGGGTTTCCCCCCTTCAGCTTGACGTCAGCGTCGA   | 500  |
| GCCGTCCAAGCAAGAAATGGAACAGTTTTTGGCGCAAAAACAGCAAGAGG  |      |
| ACGAAGAGCGAGCGCTTGCTGTACTGACCGATTTAGCGAGGGAAGAAGAA  | 600  |
| AAGGCCGCGTCTGCGCCGCCGTCCGGTCCGCTTGTCATCGGCTATCCGAT  |      |
| CCGCGACGAGGAGCCGGTGCGGCGGCTTGAAACGATCGTCGAAGAAGAGC  | 700  |
| GGCGCGTCTGTTGTGCAAGGCTATGTATTTGACGCCGAAGTGAGCGAATTA |      |
| AAAAGCGGCCGCACGCTGTTGACCATGAAAATCACAGATTACACGAACTC  | 800  |
| GATTTTAGTCAAAATGTTCTCGCGCGACAAAGAGGACGCCGAGCTTATGA  |      |
| GCGGCGTCAAAAAGGCATGTGGGTGAAAGTGCGCGGCAGCGTGCAAAAC   | 900  |
| GATACGTTTCGTCCGTGATTTGGTCATCATCGCCAACGATTTGAACGAAAT |      |
| CGCCGCAACGAACGGCAAGATACGGCGCCGGAAGGGGAAAAGAGGGTTCG  | 1000 |
| AGCTCCATTTGCATACCCCGATGAGCCAAATGGACGCGGTACCTCGGTG   |      |
| ACAAAAC TCATTGAGCAAGCGAAAAAATGGGGGCATCCGGCGATCGCCGT | 1100 |
| CACCGACCATGCCGTTGTTTCAGTCGTTTCCGGAGGCCTACAGCGCGGCGA |      |
| AAAAACACGGCATGAAGGTCATTTACGGCCTTGAGGCGAACATCGTCGAC  | 1200 |
| GATGGCGTGCCGATCGCCTACAATGAGACGCACCGCCGTCTTTTCGGAGGA |      |
| AACGTACGTCGTCTTTGACGTCGAGACGACGGGCCTGTCGGCTGTGTACA  | 1300 |
| ATACGATCATTGAGCTGGCGGCGGTGAAAGTGAAAGACGGCGAGATCATC  |      |
| GACCGATTCATGTCGTTTGCCAACCCTGGACATCCGTTGTCGGTGACAAC  | 1400 |
| GATGGAGCTGACTGGGATCACCGATGAGATGGTGAAAGACGCCCCGAAGC  |      |
| CGGACGAGGTGCTAGCCCGTTTTTGTGACTGGGCCGGCGATGCGACGCTT  | 1500 |
| GTTGCCCAACGCCAGCTTTGACATCGGTTTTTTAAACGCGGGCCTCGC    |      |
| TCGCATGGGGCGCGGCAAAATCGCGAATCCAGTCATCGATACGCTCGAGC  | 1600 |
| TGGCCCGTTTTTTTATACCCGGATTTGAAAAACCATCGGCTCAATACATTG |      |
| TGCAAAAAATTTGACATTGAATTGACGCAGCATACCGCGCCATCTACGA   | 1700 |
| CGCGGAGGCGACCGGGCATTGCTTATGCGGCTGTTGAAGGAAGCGGAAG   |      |
| AGCGCGGCATACTGTTTCATGACGAATTAAACAGCCGCACGCACAGCGAA  | 1800 |
| GCGTCCTATCGGCTTGCGCGCCCGTTCCATGTGACGCTGTTGGCGCAAAA  |      |
| CGAGACTGGATTGAAAAATTTGTTCAAGCTTGTTGTCATTGTCGCACATTC | 1900 |
| AATATTTTACCGTGTGCCGCGCATCCCGCGCTCCGTGCTCGTCAAGCAC   |      |
| CGCGACGGCCTGCTTGTCGGCTCGGGCTGCGACAAAGGAGAGCTGTTTGA  | 2000 |
| CAACTTGATCCAAAAGGCGCCGGAAGAAGTCGAAGACATCGCCCGTTTTT  |      |
| ACGATTTTCTTGAAGTGCATCCGCCGGACGTGTACAAGCCGCTCATCGAG  | 2100 |
| ATGGATTATGTGAAAGACGAAGAGATGATCAAAAACATCATCCGCAGCAT  |      |
| CGTCGCCCTTGGTGAGAAGCTTGACATCCCGGTTGTCGCCACTGGCAACG  | 2200 |

FIG. 88A



|   |      |
|---|------|
| TCCATTACTTGAACCCAGAAGATAAAATTTACCGGAAAATCTTAATCCAT  |      |
| TCGCAAGGCGGGGCGAATCCGCTCAACCGCCATGAACTGCCGGATGTATA  | 2300 |
| TTTCCGTACGACGAATGAAATGCTTGACTGCTTCTCGTTTTTtagggccgg |      |
| AAAAAGCGAAGGAAATCGTCGTTGACAACACGCAAAAAATCGCTTCGTTA  | 2400 |
| ATCGGCGATGTCAAGCCGATCAAAGATGAGCTGTATACGCCGCGCATTGA  |      |
| AGGGGCGGACGAGGAAATCAGGGAAATGAGCTACCGGCGGGCGAAGGAAA  | 2500 |
| TTTACGGCGACCCGTTGCCGAAACTTGTTGAAGAGCGGCTTGAGAAGGAG  |      |
| CTAAAAAGCATCATCGGCCATGGCTTTGCCGTCATTTATTTGATCTCGCA  | 2600 |
| CAAGCTTGTGAAAAAATCGCTCGATGACGGCTACCTTGTCGGGTCGCGCG  |      |
| GATCGGTCGGCTCGTCGTTTGTGCGGACGATGACGGAAATCACCGAGGTC  | 2700 |
| AATCCGCTGCCGCCGCATTACGTTTGCCCGAACTGCAAGCATTCCGGAGTT |      |
| CTTTAACGACGGTTCAGTCGGCTCAGGGTTTGATTTGCCGGATAAAAACT  | 2800 |
| GCCCGCGATGTGGGACGAAATACAAGAAAGACGGGCACGACATCCCGTTT  |      |
| GAGACGTTTTCTCGGCTTTAAAGGCGACAAAGTGCCGGATATCGACTTGAA | 2900 |
| CTTTTCCGGCGAATACCAGCCGCGCGCCACAACCTATACGAAAGTGCTGT  |      |
| TTGGCGAAGACAACGTCTACCGCGCCGGGACGATTGGCACGGTCGCTGAC  | 3000 |
| AAAACGGCGTACGGATTTGTCAAAGCGTATGCGAGCGACCATAACTTAGA  |      |
| GCTGCGCGGGCGGAAATCGACGGCTCGCGGCTGGCTGCACCGGGGTGAA   | 3100 |
| GCGGACGACCGGGCAGCATCCGGGCGGCATCATCGTCGTCCCGGATTATA  |      |
| TGGAAATTTACGATTTTACGCCGATTCAATATCCGGCCGATGACACGTCC  | 3200 |
| TCTGAATGGCGGACGACCCATTTGACTTCCATTTCGATCCACGACAATTT  |      |
| GTTGAAGCTCGATATTCTCGGGCACGACGATCCGACGGTCATTTCGCATGC | 3300 |
| TGCAAGATTTAAGCGGCATCGATCCGAAAACGATCCCGACCGACGACCCG  |      |
| GATGTGATGGGCATTTTACGACGACCCGAGCCGCTTGCGGTTACGCCGGA  | 3400 |
| GCAATCATGTGCAATGTCGGCACGATCGGCATTCCGGAGTTTGGCACGC   |      |
| GCTTCGTTCCGGCAAATGTTGGAAGAGACAAGGCCAAAAACGTTTTCCGAA | 3500 |
| CTCGTGCAAATTTCCGGCTTGTGCGCACGGCACCGATGTGTGGCTCGGCAA |      |
| CGCGCAAGAGCTCATTCAAACGGCACGTGTACGTTATCGGAAGTCATCG   | 3600 |
| GCTGCCGCGACGACATTATGGTCTATTTGATTTACCGCGGGCTCGAGCCG  |      |
| TCGCTCGCTTTTAAAATCATGGAATCCGTGCGCAAAGGAAAAGGCTTAAC  | 3700 |
| GCCGGAGTTTGAAGCAGAAATGCGCAAACATGACGTGCCGGAGTGGTACA  |      |
| TCGATTCATGCAAAAAAATCAAGTACATGTTCCCGAAAGCGCACGCCGCC  | 3800 |
| GCCTACGTGTTAATGGCGGTGCGCATCGCCTACTTTAAGGTGCACCATCC  |      |
| GCTTTTGTATTACGCGTCGTACTTTACGGTGCGGGCGGAGGACTTTGACC  | 3900 |
| TTGACGCCATGATCAAAGGATCACCCGCCATTTCGCAAGCGGATTGAGGAA |      |
| ATCAACGCCAAAGGCATTCAGGCGACGGCGAAAGAAAAAGCTTGCTCAC   | 4000 |
| GGTTCTTGAGGTGGCCTTAGAGATGTGCGAGCGCGGCTTTTCCTTTAAAA  |      |
| ATATCGATTTGTACCGCTCGCAGGCGACGGAATTCGTCATTGACGGCAAT  | 4100 |
| TCTCTCATTCCGCCGTTCAACGCCATTCCGGGGCTTGGGACGAACGTGGC  |      |
| GCAGGCGATCGTGCGCGCCCGCGAGGAAGGCGAGTTTTTGTGCAAGGAGG  | 4200 |
| ATTTGCAACAGCGCGGCAAATTGTGCAAAACGCTGCTCGAGTATCTAGAA  |      |
| AGCCGCGGCTGCCTTGACTCGCTTCCAGACCATAACCAGCTGTCGCTGTT  | 4300 |

T

FIG. 88B

|   |      |
|---|------|
| MVTKEQKERFLILLEQLKMTSDEWMPHFREAAIRKVVIDKEEKSWHFFYFQ |      |
| FDNVLPVHVYKTFADRLQTAFRHIAAVRHTMEVEAPRVTEADVQAYWPLC  | 100  |
| LAELQEGMSPLVDWLSRQTPPELKGKLLVVARHEAEALAIKRRFAKKIAD  |      |
| VYASFGFPPLQLDVSVEPSKQEMEQFLAQKQOEDEERALAVLTDLAREEE  | 200  |
| KAASAPPSGPLVIGYPIRDEEPVRRLETIVEEEERRVVQGYVFDAEVSEL  |      |
| KSGRTLTLTKITDYTNSILVKMFSRDKEDAEMLMSGVKKGMWVKVRGSVQN | 300  |
| DTFVRDLVIIANDLNEIAANERQDTAPEGEKRVELHLHTPMSQMDAVTSV  |      |
| TKLIEQAKKWGHPAIAVTDHAVVQSFEAYSAAKKHGMKVIYGLEANIVD   | 400  |
| DGVPIAYNETHRRLSEETYVVFVDTTGLSAVYNTIIELAALKVKDGEII   |      |
| DRFMSFANPGHPLSVTTMELTGITDEMVKDAPKPDEVLARFVDWAGDATL  | 500  |
| VAHNASFDIGFLNAGLARMGRGKIANPVIDTLELARFLYPDLKNHRLNTL  |      |
| CKKFDIELTQHHRAIYDAEATGHLLMRLLEKEAEERGILFHDELNSRTHSE | 600  |
| ASYRLARPFHVTLLAQNETGLKNLFKLVSLSHIQYFHRVPRI PRSVLVKH |      |
| RDGLLVGSGCDKGELFDNLIQKAPEEVEDIARFYDFLEVHPPDVYKPLIE  | 700  |
| MDYVKDEEMIKNIIRSIVALGEKLDIPVVATGNVHYLNPEDKIYRKILIH  |      |
| SQGGANPLNRHELDPVYFRTTNEMLDCFSFLGPEKAKEIVVDNTQKIASL  | 800  |
| IGDVKPIKDELYTPRIEGADEEIREMSYRRAKEIYGDPLPKLVEERLEKE  |      |
| LKSIIGHGFAVIYLYSHKLVKKSLLDDGYLVGSRGSGVSSFVATMTEITEV | 900  |
| NPLPPHYVCPNCKHSEFFNDGSGVSGFDLPDKNCPRCGTTYKKDGHDIPE  |      |
| ETFLGFKGDKVPDIDLNFSGEYQPRAHNYTKVLFGEDNVYRAGTIGTVAD  | 1000 |
| KTAYGFVKAYASDHNLELRGAEIDLAAGCTGVKRTTGQHPGGIIVVPDYM  |      |
| EIYDFTPIQYPADDTSSSEWRTHFDFHSIHDLNLLKLDILGHDDPTVIRML | 1100 |
| QDLSGIDPKTIPTDDPDVMGIFSSTEPLGVTPEQIMCNVGTIGIPEFGTR  |      |
| FVRQMLEETRPKTFSELVQISGLSHGTDVWLGNALQELIQNGTCTLSEVIG | 1200 |
| CRDDIMVYLIYRGLEPSLAFKIMESVRKKGKLTPEFEAEMRKHDVPEWYI  |      |
| DSCKKIKYMFPAKAAAYVLMVARIAYFKVHHPLLYYASYFTVRAEDFDL   | 1300 |
| DAMIKGSPAIRKRIEINAKGIQATAKEKSLTVLEVALEM CERGF S FKN |      |
| IDLYRSQATEFVIDGNSLIPPFNAIPGLGTNVAQAIVRAREEGEFLSKED  | 1400 |
| LQQRGKLSKTLLEYLESRGCLDSLDPDHNQLSLF                  |      |

FIG. 89